



CrossMark  
click for updates

## Review

**Cite this article:** He F, Murabito E, Westerhoff HV. 2016 Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering. *J. R. Soc. Interface* **13**: 20151046.

<http://dx.doi.org/10.1098/rsif.2015.1046>

Received: 5 December 2015

Accepted: 21 March 2016

### Subject Category:

Reviews

### Subject Areas:

bioengineering, synthetic biology, systems biology

### Keywords:

systems biology, synthetic biology, metabolic engineering, control engineering, metabolic networks, gene expression regulation

### Author for correspondence:

Hans V. Westerhoff

e-mail: [hans.westerhoff@manchester.ac.uk](mailto:hans.westerhoff@manchester.ac.uk)

# Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering

Fei He<sup>1</sup>, Ettore Murabito<sup>2</sup> and Hans V. Westerhoff<sup>2,3,4</sup>

<sup>1</sup>Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK

<sup>2</sup>The Manchester Centre for Integrative Systems Biology, Manchester Institute for Biotechnology, School for Chemical Engineering and Analytical Science, University of Manchester, Manchester M1 7DN, UK

<sup>3</sup>Department of Synthetic Systems Biology, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

<sup>4</sup>Department of Molecular Cell Physiology, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

Metabolic pathways can be engineered to maximize the synthesis of various products of interest. With the advent of computational systems biology, this endeavour is usually carried out through *in silico* theoretical studies with the aim to guide and complement further *in vitro* and *in vivo* experimental efforts. Clearly, what counts is the result *in vivo*, not only in terms of maximal productivity but also robustness against environmental perturbations. Engineering an organism towards an increased production flux, however, often compromises that robustness. In this contribution, we review and investigate how various analytical approaches used in metabolic engineering and synthetic biology are related to concepts developed by systems and control engineering. While trade-offs between production optimality and cellular robustness have already been studied diagnostically and statically, the dynamics also matter. Integration of the dynamic design aspects of control engineering with the more diagnostic aspects of metabolic, hierarchical control and regulation analysis is leading to the new, conceptual and operational framework required for the design of robust and productive dynamic pathways.

## 1. Introduction

Metabolic engineering focuses on the design, construction and optimization of metabolic pathways and their regulatory processes with the aim of improving the production of chemicals, pharmaceuticals or bio-fuels [1]. It also increasingly relates to strategic approaches towards improved metabolic diagnosis and therapies (e.g. [2]). A central issue in metabolic engineering consists of identifying the components or processes that control and limit the production flux [3]. The dependency of steady-state fluxes and species concentrations on process activities such as enzyme-catalysed reactions can be quantified and understood by metabolic control analysis (MCA) [4,5]. Hierarchical control analysis (HCA), an extension of MCA, takes transcriptional regulation and signal transduction also into account [6]. Both MCA and HCA rely on the explicit account of the rate laws governing the biochemical transformations in the pathway. Alternatively to these approaches, linear optimization methods, such as flux balance analysis (FBA) [7], can be used to predict the flux distribution that supports the maximal formation rate of the product of interest using a purely stoichiometric description of the entire genome-scale metabolic network [8,9]. The actual implementation of the optimal flux distribution can then be achieved using pathway design approaches such as gene knockout and overexpression of native or heterologous genes [10,11], or modulation (e.g. partial inhibition) of gene expression. The ensemble of these approaches is named 'systems metabolic engineering' in recent literature [12].

Two important design schemes are often employed in metabolic engineering. One uses and modulates native pathways. The other constructs non-natural pathways through the import of heterologous genes, the synthesis of new DNAs and new enzymes [13], or the provocation of new interactions. The latter is part and parcel of synthetic biology. Synthetic biology uses a forward-engineering approach to create new biological parts or networks, modifying existing biological systems [14,15]. Early synthetic biology primarily designed and implemented individual genetic modules, such as genetic toggle switches, logic gates, oscillators, cascades and biosensors (as reviewed in [16–18]). Recently, a ‘second wave’ of synthetic biology created larger functional cellular systems, including signalling and metabolic networks [19]. The challenges include the need of new design principles that enable one to manage the complexity of biology, and a deeper understanding of biological network functions at a ‘systems level’. The former includes the ambition to make the synthetic circuits adaptive and robust even if precise information of the pathway is not available [20,21]; the latter requires the assistance of systems biology and advanced reverse engineering methods [22]. Synthetic biology can advance metabolic engineering significantly by producing new enzymes, novel enzymatic activities and creating new metabolic pathways [13]. Pioneering synthetic biology applications to metabolic-regulatory systems include a re-designing of the native regulatory gene circuits for the enhanced production of lycopene and a synthetic gene-metabolic oscillator with assistance from nonlinear dynamic analysis [23,24]. Reviews and perspectives of synthetic biology and metabolic engineering focusing on their applications, areas of synergism and overlaps can be found in recent literature [25–29].

Formulation of products by and inside living organisms offers the advantages of ready amplification of production capacity, self-repair of the production process and specificity. Thereby, when engineering microbial metabolism in order to maximize productivity, it is important to maintain the cells’ functional stability (or a transition from one physiological state to another that does not compromise vitality) when they experience environmental perturbations or are subject to internal noise possibly including mutagenesis. Paradoxically, it can also be important for the cells to respond intensively to changes that are intentional in the production process such as a shift from a growth phase to a production phase. Such ‘agile robustness’ is inherent in most living organisms but aimed at improving the organism’s fitness rather than productivity. The metabolic pathways involved in productivity are typically subject to transcriptional, signal-transduction and metabolic regulation inclusive of feedback and feed-forward regulatory mechanisms (e.g. a repressor–inducer system) [9,30,31]. Some regulatory loops may serve homeostasis, whereas others may exhibit (non-linear) dynamic behaviour such as that of bi-stable switches or stable oscillations (e.g. the collective yeast glycolytic oscillator [32] and the gene-metabolic oscillator [24]).

Functionalities such as oscillations and switching are often engineered into man-made mechanical or electronic control systems. Accordingly, the disciplines that help do this, i.e. systems and control engineering, might be able to help understand (i) how the regulatory mechanisms in natural systems promote the organism’s fitness and (ii) whether the regulatory mechanisms that emanated from evolutionary

‘design’ are similar to, or better/worse than, what engineers would have designed. They may also help design new synthetic biological regulatory devices with properties that enhance the production process in the biological context [33,34], such as improving pathway robustness and stability. For example, three classical control mechanisms known in control engineering, i.e. proportional, integral and derivative control, have been identified in the regulation of energy metabolism in living organisms [35]. In addition, the closed-loop feedback control principle has been used to design both genetically encoded (RNA or protein) biosensors that help optimize and regulate heterologous pathways [36] and a sensor-regulator system that improves fatty acid-based biodiesel production [37]. When using engineering principles to construct a synthetic metabolic-regulatory system from isolated sub-parts or modules, several issues need to be taken into account. These include: (i) crosstalk and retroactivity effects when interconnecting biological sub-modules [38], (ii) stability and design constraints [39], (iii) response sensitivities and noise propagation [40], and (iv) interactions with their environment [41]. Almost all these issues can be analysed and interpreted using systems and feedback control theory, but have not fully been studied.

This work reviews and investigates how control and optimization approaches used in metabolic engineering as well as diagnostic, analytic and deductive approaches used in synthetic systems biology are related to approaches developed in control systems engineering. In §2, classical steady-state approaches employed in metabolic engineering, such as MCA, FBA, supply–demand theory and regulation analysis, are reviewed. Rather than exclusively discussing their roles in the design of pathways, which have been previously reviewed to some extent (e.g. [42,43]), the focus here is on their recent development for, and applications to, metabolic networks under both enzymatic and gene-expression regulation and the corresponding robustness analysis. In §3, the theoretical principles of control engineering are used to interpret the issues related with dynamic flux control and enzyme activation that have been analysed by a dynamic version of FBA. Metabolic and gene-expression regulation is related to both the proportional and the integral control of control engineering, and the MCA and HCA approaches of systems biology. In §4, synthetic biology-related design strategies, such as the re-engineering of gene circuits and allosteric proteins, together with issues such as modularity and crosstalk, design constraints and noise propagation, are discussed from a systems engineering perspective. A central issue in both the analysis and the engineering/design of a metabolic-regulatory network is a trade-off between optimality (e.g. maximum production, minimum transition time) and robustness (e.g. assurance of the system’s stability). This issue is investigated throughout the review.

## 2. Steady-state control and regulation analysis

### 2.1. Regulatory motifs

To adapt to and maintain growth under various intracellular and extracellular conditions, both long- and short-term regulatory mechanisms are often present in the same metabolic network. Relatively long-term regulation is based on changing the expression level of genes or on the rewiring of the gene regulatory networks. Such ‘gene-expression regulation’

may be crucial for cell survival in circumstances requiring extensive and persistent cellular alterations. Short-term regulation is based on modification of enzymatic activities through (i) protein–protein interactions and covalent modification (e.g. phosphorylation) in signal transduction or (ii) substrate, product or allosteric regulation of metabolic reactions. Such ‘metabolic regulation’ provides a fast response to acute extracellular and intracellular perturbations and helps the cell to control the metabolic and Gibbs energy balance. The distribution of different types of regulation of the metabolism of organisms such as *Escherichia coli* [44,45] and yeast [46] over metabolic, signalling and gene expression has been studied intensively. A comprehensive review [47] outlined the regulation of important metabolic tasks at different levels, including nutrient uptake, energy and amino acid metabolism, and protein synthesis, and highlighted a set of metabolites that carry out specific regulatory functions. In pathways like glycolysis, nitrogen assimilation, the citric acid cycle and inorganic ion uptake, gene-expression regulation is important for long-term adaptation to the availability of nitrogen and carbon sources. Shorter term adaptation of these same pathways also seems important as it has given rise to the evolutionary emergence of sophisticated fast cascade control mechanisms [45]. In some amino acid biosynthesis pathways, such as those for glutamate, tyrosine, tryptophan, phenylalanine, arginine and proline, both gene expression and metabolic regulation have been identified. Such co-regulation ensures fine-tuning of metabolic activities and the rapid response to over-accumulation of end products.

The dynamics of a metabolic pathway under both metabolic and gene-expression regulation can be represented by the following generic model:

$$\left. \begin{aligned} \frac{dX}{dt} &= N \cdot v(X, E) = N \cdot \text{diag}(E) \cdot f(X) \\ \frac{de_i}{dt} &= g(x_j) - k_i^d \cdot e_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (2.1)$$

Here  $N$  is the stoichiometry matrix indicating which reactions are involved in the metabolism of each metabolite  $x_i$ ,  $E = [e_1, e_2, \dots, e_n]^T$  denotes the set of concentrations of the enzymes catalysing the various ( $n$ ) reactions,  $\text{diag}(E)$  is a diagonal matrix with  $E$  along its diagonal,  $v$  is the vector of the reaction rates, and  $f(X)$  is a vector function of the concentrations of the metabolites  $X$  and kinetic parameters. Not all the enzymes are necessarily regulated through gene expression, although many are and their dynamics are modelled by the second equation. The gene-expression function  $g(\cdot)$  is assumed to depend on metabolite concentrations, e.g. on the concentration of the (pen)ultimate metabolite in an end product pathway.  $k_i^d$  is the degradation rate constant of the  $i$ th enzyme. In fast-growing organisms and for stable proteins,  $k_i^d$  mainly represents the dilution effect due to growth, but in other cases it will also depend on proteolysis.

In some metabolic networks, such as in a linear biosynthetic chain of reactions, two regulatory motifs are involved in controlling the flux. One is the end product feedback inhibition, typically acting on the first enzyme of the pathway; the other is the intermediate metabolite feed-forward activation of the downstream enzymes. The expression levels of enzymes may be controlled through transcription factors. These are themselves proteins whose functioning is inhibited or activated by the end product or an intermediate

metabolite, and whose expression may also be regulated. According to the connectivity architectures of the gene networks alone, several transcription–regulation motifs have been identified [30,48]. These include the single-input module consisting of a set of operons controlled by a single transcription factor, the dense overlapping regulation motif and the feed-forward loop (see [48] for more discussions).

Feed-forward loops have been identified also in signal-transduction networks as inferred from observed protein–protein interactions. Signal-transduction networks are home to other motifs that are absent from transcription networks, such as the diamond pattern and multi-layer motifs. Mixed-feedback loop motifs have been observed in composite networks that consist of both transcription–regulation and protein–protein interactions [49]. When metabolic regulation is added explicitly, various structures are reinforced. It is not clear why and how the regulatory motifs occur at different ‘hierarchical levels’, i.e. at the level of gene networks, signal-transduction networks and metabolic networks at the same time, although a recently developed ‘regulation analysis’ has confirmed that such hierarchical regulation does occur in real metabolic pathways [46,50,51]. This will be further discussed in §2.3.3. Such hierarchical regulation calls for a critical analysis of existing interpretations of the functionality of gene regulatory networks that were based on the exclusive analysis of transcription regulation, but this resides beyond the scope of this review. In the following subsections, classical metabolic engineering approaches to the control of enzyme expression and network flux are reviewed and their recent attempts to take complex regulatory mechanisms into account are examined.

## 2.2. Flux balance analysis

FBA identifies flux patterns satisfying the steady-state condition imposed by setting equation (2.1) to zero, while maximizing a given objective function  $Z$  to obtain the corresponding  $v_{\text{optimal}}$ . In standard FBA, the objective function is a mathematical representation of a biological process, such as biomass or adenosine triphosphate (ATP) production, that the organism is assumed or required to perform optimally. Commonly  $Z$  takes the form of a linear combination of the flux variables  $v$ . Then the standard FBA can be formulated as a linear programming (LP) problem as follows:

$$\left. \begin{aligned} \max \quad & Z = c^T \cdot v \\ \text{s.t.} \quad & N \cdot v = 0; \quad v^L \leq v \leq v^U \end{aligned} \right\} \quad (2.2)$$

where  $v^L$  and  $v^U$  are the lower and upper bounds, respectively, delimiting the range of possible values for the flux variables  $v$ , and  $c$  is the set of coefficients defining the objective function  $Z$  in terms of a linear combination of the fluxes  $v$ . The main advantage of FBA over other modelling approaches is that it only requires the stoichiometries of *all* the chemical reactions (without kinetic information  $f(X)$  in (2.1)), and these have become increasingly accessible with the generation of metabolic maps. In fact, it has become possible to generate genome-wide, hence essentially complete, versions of these maps thanks to the sequencing of the genomes of many organisms, the biochemical identification work done over many years, and intensive collaborations between laboratories (e.g. [9,52]). This makes FBA particularly suitable, and indeed widely used, in the study of the metabolic capabilities of an organism on a genome-wide

scale. Genome-scale metabolic reconstructions of more than 45 organisms are already available [31], and FBA constitutes one important tool for extracting the knowledge encoded in such biochemical networks.

FBA is used to address different relevant biological questions in different ways:

- *Boolean perturbation analysis*. *In silico* gene-deletion experiments can be performed by constraining the flux of the affected reactions to zero. FBA is then used to assess the criticality of these genes with respect to the criterion of optimality represented by the objective function  $Z$ . This kind of study is particularly useful for investigating synthetic genetic interactions [43], and for suggesting drug targets for drugs aimed at interfering with certain functionalities of the networks. Standard FBA cannot directly predict positive effects of gene knockouts on production or growth yield. However, one can infer that all steps with zero FBA flux may be knocked. Several more sophisticated, bi-level optimization algorithms [53–55] have been developed to predict gene knockouts that may improve product yield, and have been implemented to microbial strain design.
- *Optimization of medium composition*. By changing the lower and upper bounds ( $v^L$  and  $v^U$ ) of the so-called ‘exchange reactions’ (i.e. those reactions involved in the transport of chemical species from the external environment to the cell and vice versa), it is possible to simulate different growth conditions. This application of FBA allows one to assess the relative suitability of different growth media with regard to a given biological process that the organism is supposed to carry out optimally [7]. This is relevant for industrial protein production.
- *Flux variability analysis (FVA)*. FVA consists of running FBA in an iterative manner so as to identify the widest range of values that each reaction flux can take without compromising the level of optimality of the network. In other words, FVA addresses the question which alternative flux patterns would lead to the same optimal performance. This procedure is closely related to a robustness analysis in which the effect on the objective function of varying a particular reaction flux can be analysed in terms of the number of potential escape mechanisms of the network [56]. FVA has also been used to find alternative optimal solutions, i.e. flux distributions that are equally optimal with regard to a given biological process [57].

The original FBA formulation has been extended and modified to encompass a wider range of possible studies on the genome-scale metabolic properties of an organism:

- *Regulatory FBA (rFBA)*. The standard FBA formulation is enriched with regulatory information by integrating Boolean logic operators [58,59]. Transcriptional regulatory events are incorporated in a reconstructed metabolic model to constrain the space of possible network functions further. This allows one to analyse and predict the effects of transcriptional regulation on cellular metabolism at a systemic level (see [43,60] for overview), provided the transcription regulation is essentially binary (on–off). Recently, a matrix formalism for representing transcriptional regulatory networks was proposed [61]. A pseudo-stoichiometric matrix  $R$  (similar to  $N$  in (2.2)) was introduced. Its rows

represent extracellular metabolites, genes and gene products (i.e. proteins), and its columns describe the Boolean regulatory relationships. Boolean rules define which gene products are turned ‘ON’ and which ones remain ‘OFF’. This introduces additional constraints on the corresponding fluxes in (2.2).

- *Data-driven FBA*. One of the main factors that affect the reliability of an FBA prediction is the choice of an appropriate objective function  $Z$ . The selection of  $Z$  is the subject of active research [62,63]. The reliance upon a specific cell function to be optimal may introduce a bias that prevents us from grasping the true physiological state of the organism [64]. Different reformulations of FBA have been proposed to predict flux patterns based on a set of measured quantities. Shlomi *et al.* [65] integrated tissue-specific gene and protein expression data with a genome-scale reconstruction of the human metabolic network to describe the tissue specificity of human metabolism. Different integer values were assigned to different gene-expression states, so as to distinguish among highly (1), hardly (–1) and moderately (0) expressed genes. The objective function was set to minimize the differences between the activity of each reaction in the predicted flux pattern and the integer representation of the corresponding experimental gene-expression level. By minimizing such an objective function, the authors retrieved flux patterns where the reaction rates were more strongly correlated with their corresponding expression state. In Lee *et al.* [66], absolute gene-expression data generated through RNA-Seq were used to provide a more precise indication of enzymatic activity than that generated through relative expression techniques such as in Shlomi *et al.* [65], although lack of correspondence between mRNA and protein levels may compromise this approach if protein levels are not also taken into account [67]. Data-driven FBA may also use the exometabolome, i.e. restrict the exchange flux pattern to what is observed experimentally. This has led to reasonable confinement of all possible flux patterns [64].
- *Dynamic FBA (DFBA)*. One strength of FBA is that it is computationally affordable, as it only relies on the reaction stoichiometry in the metabolic network. This makes FBA suitable to make predictions on a genome-wide scale. At the same time, however, this also represents a limitation as FBA can only describe or predict steady states [12]. There are studies extending FBA to include dynamic behaviour. Mahadevan *et al.* [68] introduced DFBA which incorporates rates of change of flux constraints. DFBA was used to predict the dynamics of diauxic growth of *E. coli* on glucose and acetate. DFBA also provides a suitable framework for multi-scale metabolic modelling, where the interplay of different cell types and tissues is taken into account [69]. More recently, DFBA has been extended to metabolic networks coupled with gene expression of the corresponding enzymes, where it incorporated constraints on resource allocation [70,71].
- *FBA and nonlinearity*. In the standard formulation of FBA, it is implicit that the ATP cost of biomass is a constant, i.e. independent of the fluxes entertained by the network, although this is often not the case. Switching from fermentation to respiration implies the synthesis of more mitochondria, hence of a number of enzymes involved

in the tricarboxylic acid (TCA) cycle and respiratory chain. Consequently, the Gibbs energy cost of such a respiratory regime is higher than that of fermentation. Simeonidis *et al.* [64] proposed an iterative approach where a modified formulation of FBA was used to take into account the different ATP requirements of different metabolic regimes. Although the approach was implemented in the form of a LP problem, their iterative algorithm allowed the authors to investigate a nonlinear property of the system, where the total consumption of ATP was made dependent on the mitochondrial flux. Thereby, they could predict a shift from respiratory to fermentative metabolism known as the Crabtree effect and related to the Warburg effect.

## 2.3. Metabolic, hierarchical control analysis and regulation analysis

### 2.3.1. Metabolic and hierarchical control analysis

In a biochemical network, the dependence of a system variable (e.g. the rate of metabolic reaction or phosphorylation, the concentration of a metabolite, the magnitude of a transmembrane electric potential) on the kinetic activity parameters (e.g. the activity of an enzyme) can be quantified by the *control coefficients* as introduced in MCA [4,72]. The control coefficients indicate the steady-state change in the concentration of a metabolite  $X$  or flux  $J$  in response to a modulation of an effector (e.g. inhibitor)  $p$  which acts directly on the activity of the reaction step  $i$ . The formal definition of concentration control coefficient ( $C_i^X$ ) and flux control coefficient ( $C_i^J$ ) are given as follows:

$$\left. \begin{aligned} C_i^X &= \left( \frac{\partial \ln X}{\partial \ln a_i} \right)_{\text{ss}} = \frac{(\partial \ln X / \partial \ln p)_{\text{system steady state}}}{(\partial \ln v_i / \partial \ln p)_{\text{process } i \text{ steady state only}}} \\ \text{and } C_i^J &= \left( \frac{\partial \ln J}{\partial \ln a_i} \right)_{\text{ss}} = \frac{(\partial \ln J / \partial \ln p)_{\text{system steady state}}}{(\partial \ln v_i / \partial \ln p)_{\text{process } i \text{ steady state only}}} \end{aligned} \right\} \quad (2.3)$$

The complexity of these definitions warrants elaboration, as it carries much of the essence of the complexity of systems with interacting processes, as opposed to individual processes in isolation [73]. The concentration control coefficient  $C_i^X$  is defined as the effect of a small modulation of the activity  $a_i$  of the process  $i$  on the steady-state concentration of  $X$ .  $C_i^J$  measures the same for the control of flux  $J$ . The activity  $a_i$  is defined as any multiplier of the rate equation of process  $i$  that is fixed unless altered either virtually or experimentally; it does not depend on any of the concentrations of the metabolites or enzymes. In enzyme catalysed reactions,  $a_i$  may correspond to the catalytic rate constant  $k_{\text{cat}}$  assuming the  $k_{\text{cat}}$  of the forward and the reverse reaction should be modulated proportionally in order to preclude violations of the second law of thermodynamics. In metabolic networks where gene expression is constant, the most concrete instantiation of  $a_i$  is the total concentration of the enzyme itself. This instantiation is valid when the reaction rate of the enzyme catalysed reaction is indeed proportional to the concentration of the enzyme-catalyst and when the total enzyme concentration is not altered by the internal regulation of the system. If the proportionality does not exist, such as in the case of metabolic channelling or enzyme dimerization, the meaning of  $a_i$  should regress to the multiplier mentioned above. Alternatively, one then should use the

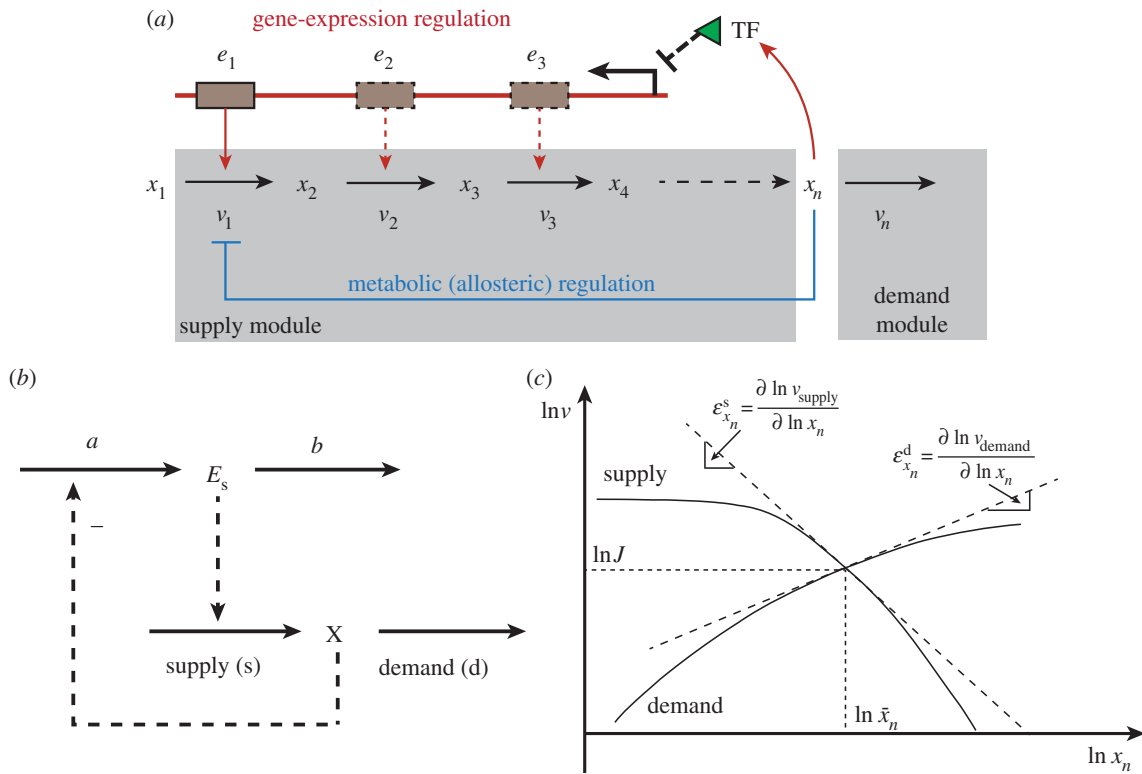
right-hand-side part of the definition, where a parameter  $p$  is used that specifically affects process  $i$ . Then the flux change between the two steady states (with and without parameter change) is compared to the effect of the parameter on the process rate  $v_i$ , had the latter been in isolation of the rest of the system but under the same nano-environmental conditions. When the parameter  $p$  is perturbed, the change in  $v_i$  in the denominator of (2.3) equals the change in the rate of reaction  $i$  only if all the other variables that affect that rate have been kept constant. This is referred to by the subscript 'process  $i$  steady-state only' and the fact that the derivative is partial: the process rates  $v_i$  are considered 'local' functions of the direct substrates, products and modifiers of the corresponding reactions  $i$ . By contrast,  $X$  and  $J$  are here considered as steady-state functions of all the parameters in the system but not of the metabolic variables.

The control coefficients describe the control exercised by a specific reaction or enzyme ('process') on the overall system variables or fluxes, while 'local' regulatory properties of individual enzymes such as  $(\partial \ln v_i / \partial \ln p)_{\text{process } i \text{ steady state only}}$  are captured by the so-called *elasticity coefficients*. Such an elasticity coefficient corresponds to the local enzyme-only response of a reaction rate to one change in its immediate environment, for example, to a change in the concentration of metabolite  $X$ . The corresponding elasticity coefficient,  $\varepsilon_X^{v_i}$ , is defined as a partial log-log derivative of  $v_i$  with respect to  $X$ .

Two important types of law have been discovered by MCA (reviewed in [67]). The summation laws restrict the sums of control coefficients over all reaction steps in the network to simple integer values indicating that the control over network properties such as fluxes, concentrations, efficiency, yields, frequencies and transient times are conserved. One implication is that removal of a flux limitation in one step will always introduce more of a flux limitation elsewhere in the network. Connectivity laws relate the global network properties that are described by control coefficients to the local elasticities, i.e. they relate network function to the collective of molecular functions. For instance, using a summation and a connectivity law, the concentration *control* coefficients for a simple two-step pathway can be expressed into their elasticities,

$$C_1^X = \frac{1}{\varepsilon_X^{v_2} - \varepsilon_X^{v_1}} = -C_2^X, \quad (2.4)$$

reflecting not only that both the upstream and the downstream steps are important for the steady-state concentration of the intermediate, but also that they are precisely equally important, and that the importance of both decreases when either step is more elastic. Similar expressions for flux control coefficients exist [67], for instance, stipulating that the more rate limiting enzymes are the ones that are least responsive to their immediate metabolic environment, hence not necessarily the ones that catalyse irreversible reactions or are first in the pathway. Classical MCA studied the control in metabolic pathways. For 'hierarchical' regulatory networks with interactions at different levels, i.e. metabolic, signal transduction and gene expression, HCA [6] was developed as a generalization of MCA. It will be discussed in the next subsection.



**Figure 1.** An unbranched metabolic pathway under hierarchical regulation. (a) The first enzyme is regulated through both transcriptional repression and allosteric activity inhibition by the end product. Enzymes in other steps might also be regulated through gene expression (in dashed arrows), but this is not explicitly considered here. TF denotes transcription factor. (b) The hierarchical supply–demand representation of the pathway (a). The lower part represents the classical metabolic supply–demand system, in which only the metabolic regulation (in this case allosteric inhibition) is considered. The letter ‘X’ denotes the penultimate product  $x_n$ , still in pathway. The supply is catalysed by enzyme  $E_s$  (i.e.  $e_1$  here or enzymes stemming from an entire operon). (c) Illustration of the steady-state properties of a supply–demand system in terms of changes in the flux, intermediate concentration and elasticity coefficients. (Online version in colour.)

### 2.3.2. Modular control analysis: hierarchical supply–demand theory

Engineering can benefit from the existence of supportive theories that are simpler to understand and are accompanied by precise theories such as MCA. To facilitate the control analysis and understanding of the functioning of large metabolic networks, modular MCA [74] has divided metabolic networks into modules with relatively autonomous activities connected through well-identified metabolites. Hofmeyr *et al.* [75,76] have developed a particular example of this, in their *supply–demand* theory. This theory serves to make the essence of the complex regulation of cell function understandable to the human. It may help design a new regulatory architecture which shifts control away from what evolved as best for the organism in its normal habitat to what is better for a metabolic engineering application. It achieves this by partitioning a pathway into supply and demand modules. More recently, supply–demand theory has been generalized from including exclusively metabolic regulation to both metabolic and gene-expression regulation, which is named the hierarchical supply–demand theory [77]. For example, an unbranched metabolic pathway under both allosteric inhibition and transcriptional regulation of the first enzyme by the end product (figure 1a) can be simplified into a hierarchical supply–demand system (figure 1b).

In such a hierarchical supply–demand system, the steady-state dependence of a system function, e.g. the concentration of an intermediate metabolite X, on the ‘overall’ change in the reaction activity of a supply (or demand)

module can be quantified by the hierarchical control coefficient employed in HCA:

$$H_s^X = \frac{\partial \ln X}{\partial \ln v_{\text{supply}}} = \frac{1}{\epsilon_X^s - {}^* \epsilon_X^s}. \quad (2.5)$$

Here,  ${}^* \epsilon_X^s$  is an ‘overall’ elasticity coefficient [72], including a classical ‘direct elasticity’ only related with metabolic responses and an ‘indirect elasticity’ due to gene-expression regulation, i.e.  ${}^* \epsilon_X^s = \epsilon_X^s + \epsilon_{E_s}^s \cdot c_a^{E_s} \cdot \epsilon_X^a$ . In HCA, the lower case  $c$  is used for metabolic control coefficients in a local network (i.e. metabolic or gene expression but not their combination). The control of gene expression  $c_a^{E_s} = 1/(\epsilon_{E_s}^b - \epsilon_{E_s}^a)$ .  $\epsilon_{E_s}^s$  is often equal to 1, i.e. when the rate of the reaction in isolation (e.g. supply module) is proportional to the concentration  $E_s$  of the enzyme catalysing it. As a result, the hierarchical concentration control coefficient (2.5) can be further expressed in terms of all the elasticity coefficients in the network:

$$\begin{aligned} H_s^X &= \frac{1}{\epsilon_X^d - \epsilon_X^s - ((\epsilon_{E_s}^s \cdot \epsilon_X^a)/(\epsilon_{E_s}^b - \epsilon_{E_s}^a))} \\ &= \frac{1}{\epsilon_X^d + (-\epsilon_X^s) + ((\epsilon_{E_s}^s \cdot (-\epsilon_X^a))/(\epsilon_{E_s}^b + (-\epsilon_{E_s}^a)))} \\ &= -H_d^X. \end{aligned} \quad (2.6)$$

The terms in parentheses are usually positive. The equation shows that the control by supply (i) decreases with the absolute magnitudes of the elasticities of the supply, demand and protein synthesis with respect to X, but (ii) increases for increasing elasticities of the protein synthesis

and degradation reactions with respect to the concentration of the enzyme ( $E_s$ ).

For the unbranched regulatory pathway given in figure 1a, only the first enzyme is regulated by the end product  $x_n$  through metabolic and transcriptional regulation. Hence, the 'total elasticity' in the supply module becomes  $*\varepsilon_x^s = *\varepsilon_{x_n}^s = \varepsilon_{x_n}^s + \varepsilon_{E_1}^{v_1} \cdot c_{Trans1}^{E_1} \cdot \varepsilon_{x_n}^{Trans1}$ , and the total elasticity through the metabolic regulation depends on the first and the  $(n-1)$ th step as  $\varepsilon_{x_n}^s = c_1^{f_1} \cdot \varepsilon_{x_n}^{v_1} + c_{n-1}^{f_{n-1}} \cdot \varepsilon_{x_n}^{v_{n-1}}$ . Because both metabolic and gene-expression regulation constitute negative feedbacks and  $x_n$  inhibits the activity of  $(n-1)$ th step, the total elasticity in the supply  $*\varepsilon_{x_n}^s$  is a monotonically decreasing function of  $x_n$  and becomes more negative with increasing  $x_n$ .  $\varepsilon_{x_n}^d = \varepsilon_{x_n}^{v_n}$  is an increasing function of  $x_n$  but decreases asymptotically to zero with increasing  $x_n$ . The intersection of these two monotonic functions determines the steady-state properties of the pathway, i.e. the steady-state concentration of metabolite, the steady-state flux, the elasticities with respect to the supply and demand modules as shown in figure 1c, and thereby the distribution of control between supply and demand [75–77].

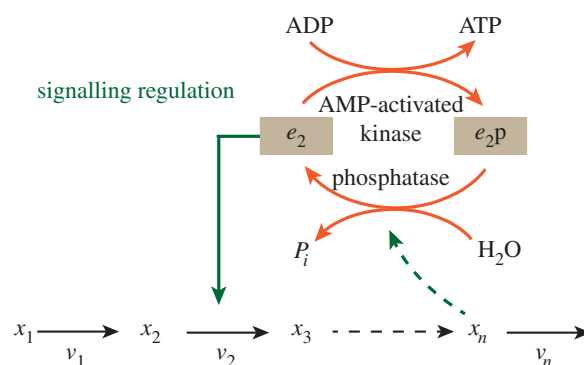
Some recent studies have extended standard MCA to take account of network uncertainties [78], to improve sampling of elasticity parameters using thermodynamic information [79] and to construct genome-scale MCA guided networks [80]. The appreciation that metabolic control does not reside in metabolism alone but also in transcription and translation, and in fact in all these, and other [81–83].

### 2.3.3. Regulation analysis: the diverse organism's response to perturbation

The magnitude of a flux or concentration 'control' coefficient employed in MCA or HCA represents a potential effect on the flux or metabolite concentration if the activity of a reaction step is modulated. This magnitude however does not indicate whether an enzyme or a kinase of a specific reaction step is actually activated by the network components (i.e. self-regulation) or modulated by an experimenter in a synthetic design [84], or not modulated at all.

To describe such regulatory information, Sauro [85] proposed a so-called 'partitioned regulatory coefficient' which describes how a perturbation in the rate of reaction  $i$  affects the enzyme activity of another reaction step  $j$  in terms of the change in the flux  $J_j$  and the altered concentrations of the metabolites that interact directly with that enzyme. This analysis only addressed regulation at the metabolic level such as allosteric or competitive effects. Ter Kuile & Westerhoff [86] generalized this and developed a hierarchical 'regulation analysis' to accommodate regulation both by metabolic effects of chemical conversions and by 'hierarchical' effects of information transfer that includes signalling or gene expression. This approach has been successfully used to analyse the steady-state regulatory properties of several important metabolic pathways [46,50,86,87] typically showing that regulation tends to be distributed across all levels in the regulatory hierarchy.

More specifically, the rate  $v_i$  of an enzyme-catalysed reaction typically depends linearly on two functions which we here name  $h$  and  $f$ . The former is related to hierarchical effects where the changes in the rate of the enzyme are due to changes in enzyme concentration or covalent modification. The latter is related to metabolic effects where changes in



**Figure 2.** An illustration of signal-transduction regulation of a metabolic pathway. Signal-transduction regulation can be any post-translational (covalent) modification, such as phosphorylation or acetylation, of an enzyme or transcription factor that is an output of a signal-transduction pathway. In this example, the enzyme that catalyses the second reaction of a metabolic pathway is subject to covalent modification through phosphorylation by an adenosine monophosphate (AMP)-dependent protein kinase.  $e_2$  denotes the fraction of the enzyme that is in the active state and  $e_{2p}$  denotes the fraction of enzyme that is phosphorylated and hence inactive. Another example is the covalent modification of glutamine synthase where AMP groups are covalently attached to the enzyme by an adenylyl transferase [88,89]. It may be noted that intermediates of central metabolism such as acetyl phosphate or the glycolytic intermediate 1,3-bisphosphoglycerate can directly cause covalent modifications of lysine residues of enzymes [90,91]. This then is metabolic regulation of metabolism through covalent modification, and constitutes yet another mechanism, which does not depend on signal transduction. (Online version in colour.)

rate are caused by changes in the concentrations of substrates, products and metabolic effectors. By denoting the rate  $v_i$  at steady state by the flux through the enzyme  $J$  (where  $v_i$  is the property of a single process  $i$ , while the flux  $J$  is a collective property and hence equal for a number of reactions in the same pathway, possibly including the process  $i$ ), we realize that

$$J = v_i = h(a_i) = h(e_i, \varphi_{a,i}) \cdot f_i(X). \quad (2.7)$$

Here  $e_i$  represents the concentration of the enzyme catalysing the process  $v_i$ . In the case of covalent modification,  $h(a_i)$  can be expressed as  $h(e_i, \varphi_{a,i}) = e_i \cdot \varphi_{a,i}$  with  $\varphi_{a,i}$  denoting the fraction of the enzyme that is in the covalent modification state that is active catalytically (figure 2). As a result, the change in the logarithm of the steady-state flux  $J$  can be expressed as

$$\left. \begin{aligned} d \ln J &= d \ln h(e_i) + d \ln f_i(X) \\ &= d \ln e_i + d \ln \varphi_{a,i} + d \ln f_i(X) \\ \Rightarrow 1 &= \frac{d \ln h(e_i)}{d \ln J} + \frac{d \ln f_i(X)}{d \ln J} = \rho_h^i + \rho_m^i \\ \Rightarrow \rho_h^i &= \frac{d \ln h(e_i)}{d \ln J} = \frac{d \ln e_i}{d \ln J} + \frac{d \ln \varphi_{a,i}}{d \ln J} = \rho_g^i + \rho_s^i. \end{aligned} \right\} \quad (2.8)$$

Here  $d$  may seem to refer small changes or perturbations, but this is not a necessary limitation in regulation analysis to the extent that an enzyme rate is directly proportional to the enzyme concentration [69]. The hierarchical regulation coefficient  $\rho_h^i$  comprises both gene-expression  $\rho_g^i$ , and signal-transduction regulation  $\rho_s^i$ . The sum of gene-expression, signal-transduction and metabolic regulation is always the

same and equal to 1 [86]. This hierarchical regulation analysis has also been extended to time-dependent cases [92,93].

Importantly, the hierarchical regulation coefficient is typically not equal to the inverse of the flux control coefficient  $C_i^f$  defined in MCA (equation (2.3)). This is because commonly the living system changes the concentrations of many enzymes at the same time, and the flux can be affected by several reactions or enzymes [67]. We shall consider the example of a two-step pathway (catalysed by enzymes  $e_1$  and  $e_2$ ) with intermediate  $X$  in which the cell responds to an external challenge by changing the relative activities of the two genes encoding the two enzymes respectively (i.e. through  $d \ln g_1$  and  $d \ln g_2$ ). The gene-expression regulation of the first enzyme is

$$\rho_g^1 = \frac{C_{g_1}^{e_1} \cdot d \ln g_1}{C_1^f \cdot C_{g_1}^{e_1} \cdot d \ln g_1 + C_2^f \cdot C_{g_2}^{e_2} \cdot d \ln g_2}. \quad (2.9)$$

$C_{g_i}^{e_i}$  denotes the control of gene  $i$  on the steady-state concentration of enzyme  $i$ . This equation indicates that the regulation coefficient is not simply the inverse of the corresponding flux control coefficient  $C_i^f$ , unless the corresponding enzyme has a flux control close to 1 (i.e.  $C_1^f = 1$  and  $C_2^f = 0$ ), or the enzyme is the sole enzyme in the pathway that is regulated by gene expression (i.e.  $d \ln g_2 = 0$ ). It has been proved [94] that there exists a relationship between the metabolic control coefficients  $C_i^f$  and gene-expression regulation coefficients  $\rho_g^i$  in an unbranched pathway, and such a relationship can be generalized to hierarchical regulation coefficients  $\rho_h^i$  by further including the effects stemming from signal-transduction regulation:

$$\sum_{i=1}^n C_i^f \cdot \rho_h^i = \sum_{i=1}^n C_i^f \cdot (\rho_g^i + \rho_s^i) \equiv 1. \quad (2.10)$$

This law implies that (i) if there is only a single rate limiting step in an unbranched pathway and that step is being regulated, the hierarchical regulation coefficient of that enzyme is always 1, (ii) if flux control is distributed and only one reaction step is regulated hierarchically, its regulation coefficient equals the inverse of the control coefficient, i.e. there is a strong hierarchical regulation if the regulated step has little flux control, and (iii) if all regulation coefficients are equal they must equal 1.

Preliminary analyses show that the above law also applies to branched pathways, with a more complex interpretation however of the regulation coefficients. Indeed, there is nothing inherent in MCA, HCA or regulation analysis that limits these approaches to linear pathways, although most examples in the literature have been simple and thereby linear. A hierarchical control and regulation analysis of cyclic networks such as the TCA cycle should be of genuine interest. After early pioneering work by Wright and co-workers [95], this is now being taken up again [96]. A case in point is the TCA cycle where much of intermediary metabolism is controlled by and sometimes also regulated through the combined levels of all TCA-cycle intermediates.

### 2.3.4. Robustness and fragility

Organisms are often subject to changes in their internal and external environments. The robustness against such perturbations is a determinant of the fitness of living organisms. Because living systems depend fundamentally on

non-equilibrium processes [72,84] there is no guarantee that they ultimately return to their original state. There are various aspects to robustness, which warrants multiple definitions. There is the maintenance of a *specific* functionality or variable of the system, where it is not required that all the variables of the system remain unchanged, i.e. homeostasis. This is known as disturbance rejection in control engineering. And, there is the transition to a new state after which the system is maintained in the new conditions. This is called robust tracking in the control engineering context. Finally, previous studies have investigated the effect of deleting one or multiple steps (e.g. knockout mutations or enzyme deficiencies) in metabolic pathways, which is termed 'structural robustness' [97,98]. A physiologically more realistic scenario may be the adaptation of the network when a reaction step is only partially inhibited, which is known as 'dynamic robustness' [99].

In a metabolic network, if the overall flux is robust to perturbations with respect to a specific step, this implies that this step is not the rate limiting step but on the contrary exerts little flux control. This suggests that there is a relationship between dynamic robustness and MCA. Quinton-Tulloch *et al.* [100] defined the robustness as the ratio between a 1% sustained perturbation in a process (i.e. perturbing the activity of an enzyme) and the corresponding percentage change in the steady-state biological function (e.g. flux). Mathematically, such a so-called MCA-based robustness coefficient equals the inverse of the corresponding control coefficient of that step. He *et al.* [77] studied the robustness of a metabolic network subjected to both metabolic and gene-expression regulation, and generalized the MCA-based analysis to HCA-based robustness analysis. The MCA (or HCA)-based robustness and fragility coefficient can be defined as follows:

$$\left. \begin{aligned} \mathcal{R}_i^f &\equiv \frac{1}{(\partial \ln f / \partial \ln a_i)} \equiv \frac{1}{C_i^f} \text{ or } \frac{1}{H_i^f} \\ \text{and} \quad F_i^f &\equiv \frac{1}{\mathcal{R}_i^f} \equiv \frac{\partial \ln f}{\partial \ln a_i} \equiv C_i^f \text{ or } H_i^f. \end{aligned} \right\} \quad (2.11)$$

Here  $f$  denotes the system function of interest, such as a particular flux  $J_j$  or a metabolite concentration  $x_j$ . Quinton-Tulloch *et al.* [100] calculated the flux robustness coefficients (FRCs)  $\mathcal{R}_i^f$  in realistic models of the glycolytic pathway of *Trypanosoma brucei* and of other *in silico* pathway models. They found that for the vast majority of the steps the individual FRCs were much larger than 1, reflecting small flux controls in those enzymatic steps. An increase in the robustness with respect to one step (e.g. glucose transporter) might come at the cost of a decrease in robustness with respect to other steps. However, the total robustness is not conserved under such operations, i.e. the sum of all the FRCs is not a constant. Unlike robustness, the MCA (or HCA)-based fragility is a conserved property because of the summation law (see §2.3.1).

Helping to understand the distribution of the control over the processes in a metabolic pathway, MCA has important applications for improving metabolite production through flux control [42] and for the identification of drug targets [101–103]. In addition, an elaborate example of using MCA and HCA to study hierarchical regulation and robustness has been reported [104]. A homeostatic control of DNA supercoiling in *E. coli* is conferred not just by enzyme activity



(metabolic regulation) or just by enzyme expression (gene-expression regulation) but an unequal mixture of both. The former is responsible for 70% and the latter for 30% of the homeostasis. And even within gene-expression regulation there is a division between transcription and translation. Such control and regulatory information obtained from MCA or HCA can be crucial for engineering genetic circuits in a metabolic pathway by indicating not only which genes (or enzymes) should be engineered, but also at which level (e.g. transcription or translation) the gene circuits should be manipulated.

Although most MCA and much HCA have been developed in the context of pathways operating at steady state, it is not confined to steady states. Indeed, MCA has been extended [67,105–107] and applied to periodic phenomena such as glycolytic oscillations [108,109] and the cell cycle [110], to transient phenomena such as transient phosphorylation of signal-transduction proteins [111] and even to bistability [112]. It is especially interesting to consider such dynamic generalizations in a control systems context [113,114]. For a metabolic-regulatory network under both metabolic and transcriptional regulation, one can study the dynamic enzymatic control in terms of a control engineering framework and then ask how this then links in with MCA and HCA. These questions will be addressed in the next section.

### 3. Dynamic control and control-theoretic analysis

A central task in metabolic engineering is the optimization of pathway flux so as to maximize the metabolite or biomass production. The analytic approaches discussed in the previous section focused on establishing which molecular properties in a network determine the network's performance. Once this is known, up- or downregulations of genes that encode steps with much control may be used to improve steady-state pathway fluxes. One may overexpress enzymes that reside on the optimal paths calculated through FBA or overexpress the enzymes with the highest control coefficients over the requested flux or yield as computed by using MCA [115,116]. These strategies can maximize pathway fluxes but in the meantime they may also decrease cellular growth rate due to the reduced expression of enzymes with high control on that growth rate, e.g. through a protein burden effect [117]. 'Dynamic metabolic engineering' or 'dynamic control' strategies [118,119] have been proposed to control gene expression and enzyme activity dynamically such as to allow trade-offs between growth and production. Models that predict the impact of genetic modifications on growth rate were developed recently [120,121]. From a systems engineering viewpoint, this corresponds to a class of optimal control problems [122]. Here, time-dependent metabolite concentrations or enzyme activities cannot always be ignored. Dynamic metabolic control also relies on advances in synthetic biology to create genetic sensors and actuators, which will be further discussed in the sections that follow.

Apart from maximizing the metabolic production rate or yield, it is important to ensure pathway homeostasis and robust adaptation to perturbations especially for the effective implementation of a 'dynamic control' strategy. This is also true for engineering systems design where a trade-off between the optimal control performance and the robustness of the designed control system needs to be achieved. Robust

perfect adaptation in a bacterial chemotaxis signalling system, in mammalian iron and calcium homeostasis, and in yeast osmoregulation have been interpreted as integral feedback control systems [123–125]. This suggests that engineering approaches may have some bearing on biological systems. One may therefore ask whether, for a genome-scale metabolic-regulatory system, robust adaptation at various hierarchical levels (i.e. metabolic and gene expression) such as observed in biology as a result of evolution, can be related to the different optimal control strategies designed in the discipline of control engineering. Does Nature beat human design? And, how could the application of control engineering to existing biological networks enhance those networks in the metabolic engineering or even the therapeutic sense? How could this lead to a better synthetic biology? Recent progress along these lines is discussed in this section.

#### 3.1. Dynamic flux balance analysis and optimal control

FBA is mainly used to study the metabolic flux distribution at steady states; the dynamics of the pathway is rarely considered. However, the actual metabolite concentrations and enzymatic activities may be time-dependent and their temporal distribution could significantly affect the performance of the pathway. In addition, the ability to re-programme the optimized pathway and to reproduce features of the microbial growth process may be important for an engineering intervention that is minimally invasive.

DFBA (as reviewed in §2.2) extends standard FBA (2.2) by including the dynamics of the metabolic networks in the optimization process:

$$\left. \begin{array}{l} \max_{X(t), v(t)} \quad Z = f(X(t), v(t)) \\ \text{s. t.} \quad \frac{dX(t)}{dt} = N \cdot v(t) \\ v^L \leq v(t) \leq v^U \\ X(0) = X_0, \quad v(0) = v_0. \end{array} \right\} \quad (3.1)$$

This is very similar to an optimal control problem studied in control engineering. The metabolic concentrations  $X$  and process rates  $v$  can be interpreted as the state variables (state functions in thermodynamics) [72,86] and the process activities  $e_i$  (see (2.1)) as control variables or inputs. If the regulation of reaction rates is explicitly considered, for example, through allosteric control from metabolites, the optimization of the overall regulatory network corresponds to a state-feedback control problem. Under certain assumptions and when a quadratic objective function is employed, e.g. transitional responses in homeostasis, this DFBA can be formulated into a classical linear quadratic regulator (LQR) problem, i.e. DFBA–LQR [126], and solved efficiently using linear control theory. Interestingly, a connection between DFBA and MCA is identified here, where the flux control coefficients in MCA are shown to be correlated with the optimal gain computed from DFBA–LQR in a feedback regulatory network.

In the above paragraph, we used the word 'control' in the sense in which it is used in control engineering. This sense differs from the connotation of the same word in MCA that we used in earlier sections. The 'control' in MCA or HCA refers to the control exercised by a time-consuming [67] process on the system's performance. It is quantified as the

sensitivity of that performance (e.g. a flux, a concentration or a cycle time) to changes in the process activity. It is called 'control coefficient' to distinguish it from the more general 'sensitivity', where the causative event that produces the change in the performance is a change in any parameter, such temperature or an equilibrium constant in the system. The 'control' in control engineering can be understood as a mechanism that could be or has been put in place by an engineer so as to improve the performance of the system. It corresponds to a strategy to alter the performance of a system. A feedback control loop in control engineering is analogous to a feedback regulatory mechanism (e.g. allosteric or transcriptional regulation) identified in a biological network and called regulatory strength in systems biology [88,127,128]. The usual aim of control engineering is to design a 'controller' that can improve the system's performance in tracking a signal or reject a disturbance to ensure robustness. Hence, the 'control' in control engineering is a strategy and in MCA it is a fact. One of the interests we have in the interface between MCA and control engineering is the use of MCA to determine which parameters are controlling most the properties of interest and to what extent, and then the use of control engineering to design extra networks and extra network properties so as to improve the robust behaviour of the biological system.

A sub-problem of DFBA is the dynamic optimal enzyme activation in biosynthetic pathways that has been studied by several authors. Klipp *et al.* [129] showed sequential enzyme (or gene expression) profiles that minimize the transition time needed to convert the substrate into the product and the sequence of which matched the enzyme order in the unbranched metabolic pathway. This 'just-in-time' activation profile was in agreement with experimental findings in the amino acid biosynthesis pathway of *E. coli* [130]. The analysis of a leucine biosynthesis pathway in yeast [131] indicated a more complex behaviour in which groups of enzymes are expressed either quickly at low amplitudes or slowly at high amplitudes. The influences of protein abundance and protein synthesis capacity on the optimal activation strategy have also been investigated [132]. A previously reported sequential activation strategy was only optimal if protein abundance relative to protein synthesis capacity was high; as protein abundance decreased, the strategy shifted to the simultaneous activation of all enzymes. These numerical approaches can be formulated into a more rigorous theoretical framework using optimal control theory [133,134] and by formulating the following optimization problem:

$$\left. \begin{array}{l} \min_{U(t), t_f} J(X(t), U(t)) \\ \text{or } \max_{U(t), \mu(t)} J(X(t), \mu(t)) \\ \text{s. t. } \frac{dX(t)}{dt} = N \cdot v(X(t), U(t)) \\ h(X(t), U(t)) \leq 0 \\ U^L \leq U(t) \leq U^H. \end{array} \right\} \quad (3.2)$$

The objective function represents either a minimization of the transition time to reach a given state of the system (or enzyme cost) or a maximization of biomass limited by for example the synthesis of the product of the pathway.  $U(t)$  is the vector of independent control variables representing enzyme concentrations  $E(t)$ ;  $\mu(t)$  is the growth rate profile. The constraints

are the system dynamics, the inequality path constraints  $h$  that restrict the synthesis capacity of individual enzymes and total amount of metabolites, and the lower and upper bounds for the control variables, i.e. the minimum and maximum amount of enzyme and the  $k_{\text{cat}}$  available for each reaction.

This optimal control problem can be solved using indirect methods such as Pontryagin's maximum principle or as a non-linear programming problem. Recently, this optimal control formulation has been further generalized to multi-objective optimization, such as a trade-off between effectiveness (e.g. transition time) and economy (e.g. enzyme consumption), and for more realistic branched pathways [135]. Global optimization methods are required to solve such a more general problem.

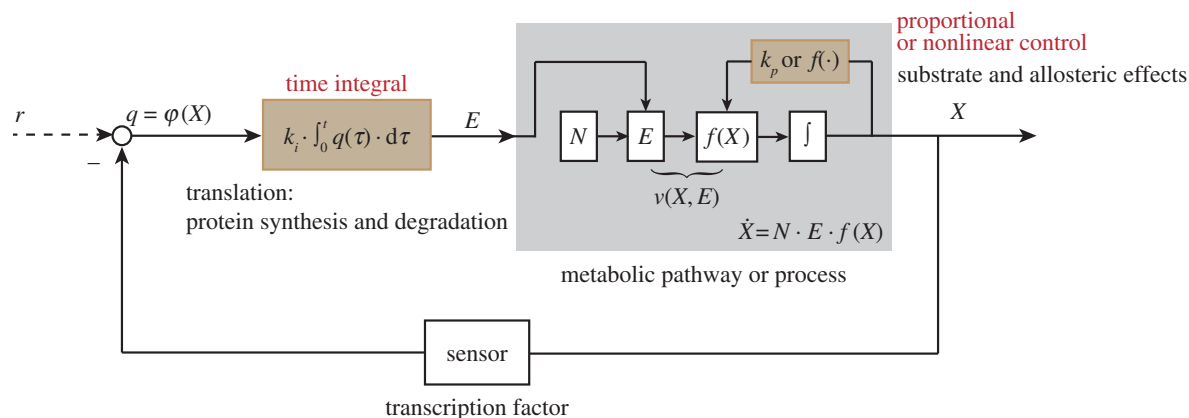
### 3.2. Metabolic and gene-expression regulation: proportional and quasi-integral control

The optimization methods reviewed in the previous subsection are mainly used to obtain an optimal flux distribution or enzyme profile. They are less concerned with the underlying control mechanism even if applied to a feedback regulatory network. In reality, the control of the amount of enzymes through gene-expression (transcriptional) regulation and the control of enzymes' catalytic behaviour through metabolic (allosteric) regulation correspond to two types of feedback control mechanisms operating at different time scales, as shown in figure 3. This may ensure the robustness versus perturbations at various frequencies as widely considered in engineering system design.

It has been shown [35,77,123] that metabolic regulation through allosteric or more direct substrate-product effects is related to a 'proportional or nonlinear control' action, because the catalytic activity of an enzyme can depend on a metabolite concentration in a proportional or nonlinear kinetic relationship. The metabolic regulation often acts as a fast actuator/controller that rapidly buffers against high-frequency perturbations but possibly with small amplitude or capability, such as adaptation to small fluctuations in the flux demand [86]. When such a fast actuator is saturated, indicating that the activity of that enzyme may be approaching its maximum capacity ( $V_{\text{max}}$ ), the system has a second 'actuator' through gene expression which is slow but leads to increase (or decrease) in the concentration of that enzyme [86]. The gene-expression regulation is slow but may be able to accommodate very large and persistent perturbations and was identified to be related to a (quasi-) integral control action [77].

Recent analyses indicate that gene-expression regulation improves the robustness of a metabolic pathway significantly but that in practice such robustness may not be infinite, because perfect adaptation requires biochemically unrealistic features such as zero-order kinetics of protein degradation [77,136]. Such control engineering interpretations can also be linked with classical MCA and HCA. The relatively fast metabolic regulation (proportional control) is related to the 'direct elasticities' of MCA, while the slow gene-expression regulation (integral control) corresponds to the 'indirect elasticities' of HCA.

Let us take ATP energy metabolism as an example. This important intracellular process can be abstracted by a supply-demand system as illustrated in figure 4. The supply (s) process represents catabolism (e.g. glycolysis) that breaks



**Figure 3.** The closed-loop control structure of a metabolic network subject to metabolic and gene-expression regulation. The substrate or allosteric effects correspond to proportional or nonlinear controls that are often modelled together with the dynamics of the metabolic process. The gene-expression regulation can be described as a (quasi-) integral control, since protein synthesis requires time integration. The transcription process or the transcription factor can be regarded as a sensor of the control system. The metabolite concentration  $X$  is the output of the system and is the controlled dependent variable. Reaction rates  $v(X, E)$  are functions of both enzyme concentration  $E$  and metabolic process  $f(x)$ . Here, the integral control input  $q$  is a function of the controlled dependent variable and a reference signal  $r$ .  $q$  compares the level of  $X$  with the level of  $r$  and differences between the two are time integrated and translated into alterations into altered enzyme levels, which then feed into the metabolic pathway. (Online version in colour.)

down nutrients and produces ATP (from adenosine diphosphate (ADP) and phosphate). The demand (d) process represents anabolism that constructs macromolecules and consumes ATP. The concentration of ATP (i.e.  $[ATP]$ ) equals  $C - [ADP]$ , with the moiety conservation sum  $C$  a constant here because the catabolic/anabolic reactions only convert ATP into ADP or vice versa. The supply is catalysed by enzyme  $E$  that denotes enzymes of catabolic reactions which are here assumed to be encoded in a single operon or regulon. The gene expression of the enzyme(s)  $E$  is increased in proportion to the concentration of ADP. The metabolic regulation addresses the interplay between the supply and demand processes.

The dynamics of ADP and enzyme  $E$  will here be described by simple kinetics

$$\left. \begin{aligned} \frac{d[ADP]}{dt} &= -k_s \cdot E \cdot [ADP] + k_d \cdot (C - [ADP]) \\ \text{and } \frac{dE}{dt} &= k_a \cdot [ADP] - k_b \cdot E - k_0, \end{aligned} \right\} \quad (3.3)$$

where the degradation of  $E$  is assumed to be a mixture of the zero- and first-order processes.  $k_0$  and  $k_b$  are the zero- and first-order protein degradation rate constants.  $k_a$  is the protein synthesis rate constant.  $k_s$  and  $k_d$  are the rate constants related to the supply and demand processes. The closed-loop control system structure of the pathway is shown in figure 5.

In this control system, the ADP concentration is a controlled (output) variable and the enzyme concentration  $E$  a manipulated (input) variable in the gene-expression feedback control loop. The zero-order degradation rate  $k_0$  can be treated as a reference signal to the system. The metabolic regulation is included as part of the ADP kinetic process. By considering a perturbation of the demand process (i.e.  $\delta k_d$ ) and reformulating the kinetics of ADP and  $E$ , one obtains

$$\begin{aligned} \delta[ADP] &= -(k_s \cdot E_{ss} + k_d) \cdot \delta[ADP] \\ &\quad - k_s \cdot [ADP]_{ss} \cdot \int_0^\infty (k_a \cdot \delta[ADP] - k_b \cdot \delta E) \cdot dt \\ &\quad + (C - [ADP]_{ss}) \cdot \delta k_d, \end{aligned} \quad (3.4)$$

where the subscript  $ss$  denotes the steady-state value and the left-hand side of the equation is a perturbation in the time

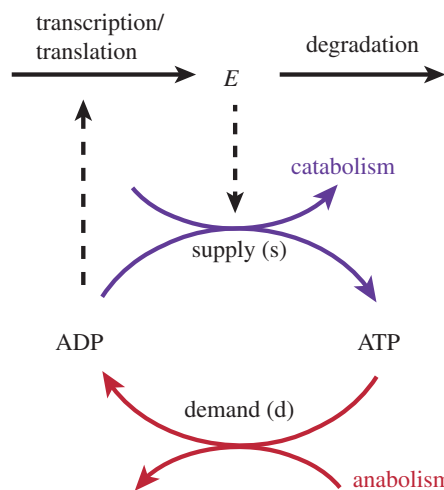
dependence of the ADP concentration. On the right-hand side, the first term is a proportional response term, the second an integral response term and the third the perturbation term. The proportional response corresponds to the direct 'elasticity' of the supply and demand reactions with respect to ADP, which is a metabolic and instantaneous regulation. The integral response is related to the protein synthesis and degradation and thus to the gene-expression regulation. If  $k_b = 0$ , the second term corresponds to an ideal integral action. By further removing the time dependence of the change in ADP using the steady-state conditions [77], the hierarchical control coefficients, quantifying the control of the enzyme level and the flux by the demand reaction, are obtained:

$$\left. \begin{aligned} H_{k_d}^E &= \frac{\delta \ln E}{\delta \ln k_d} = 1 - \frac{1}{1 + (k_s \cdot ([ADP]_{ss})^2 / (k_d \cdot C)) \cdot (k_a / k_b)} \\ \text{and } H_{k_d}^J &= \frac{\delta \ln J}{\delta \ln k_d} = 1 - \frac{[ADP]_{ss} / C}{1 + (k_s \cdot ([ADP]_{ss})^2 / (k_d \cdot C)) \cdot (k_a / k_b)}. \end{aligned} \right\} \quad (3.5)$$

Both the control of enzyme level and of demand flux are equal to 1 minus a hyperbolic function of  $k_b$ . For an ideal integral control scenario with  $k_b = 0$ , the enzyme concentration  $E$  perfectly tracks the activity of the pathway degrading ATP, and  $H_{k_d}^E = 1$ . More importantly, the pathway flux perfectly tracks the perturbation in the demand flux and  $H_{k_d}^J = 1$ . The control of  $k_d$  on ADP is zero and thereby this is the case of robust perfect adaptation. In practice, however, the dilution or the first-order proteolysis effect cannot be ignored and often  $k_b \neq 0$ . In such cases, the adaptation of the pathway to the perturbation is not perfect. An MCA-based robustness coefficient has been proposed by Quinton-Tulloch *et al.* [100] and can be expressed in terms of the various parameters:

$$\begin{aligned} \mathfrak{R}_{k_d}^{ADP} &= \frac{1}{\partial \ln [ADP]_{ss} / \partial \ln k_d} \\ &= \frac{k_d \cdot C + k_s \cdot ([ADP]_{ss})^2 \cdot (k_a / k_b)}{(C - [ADP]_{ss}) \cdot k_d}. \end{aligned} \quad (3.6)$$

Only when  $k_b = 0$  does the pathway exhibit infinite robustness ( $\mathfrak{R}_{k_d}^{ADP} = \infty$ ) to the external or parametric perturbation.



**Figure 4.** Illustration of ATP energy metabolism. This ATP supply–demand system is under both gene-expression and metabolic regulation. (Online version in colour.)

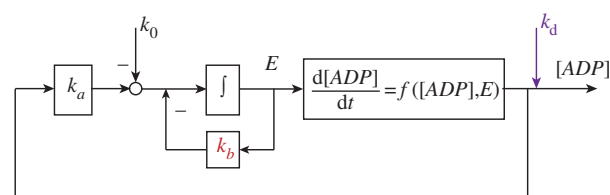
This example shows the consistency of control engineering and classical metabolic and HCA in understanding the adaptation of a metabolic pathway under both gene-expression and metabolic regulation. It also shows that a control engineering concept such as perfect adaptation can serve as inspiration for the understanding of the evolutionary design of the regulation of metabolic pathways.

## 4. Synthetic biology: design of metabolic-regulatory systems

To achieve optimal control of chemical production and to maintain pathway robustness in response to environmental changes, a number of elements in gene circuits can be designed and tuned at the transcription and translation stages. Post-translational control can be achieved via the modification of native allostery and the design of scaffold proteins. These synthetic biology-related design problems together with other issues related to the design, such as modularity, retroactivity, noise and stochastic effects, are reviewed in this section. Here, we focus on the more novel design of pathways at system level characteristic of synthetic biology, although most studies have considered the individual protein level, such as when modifying the protein sequence/function, activity or when designing new proteins in protein engineering [13].

### 4.1. Engineering gene expression and design constraints

Several important elements (or parameters) in gene circuits, such as promoters for controlling transcription initiation rate, ribosome binding sites (RBS), riboswitches controlling translation levels, plasmid replicons controlling gene copy number, small inhibitory RNAs and long non-coding RNAs, can be tuned and manipulated with the aim of achieving a desired control over metabolism [16,137]. Pioneered implementations include a re-design of the native regulatory gene circuits for the enhanced production of lycopene and a synthetic gene-metabolic oscillator [23,24]. Since simple over-expression of flux-controlling enzymes can be toxic and can slow down cell growth in understandable ways [117], synthetic gene-metabolic feedback circuits have been built to



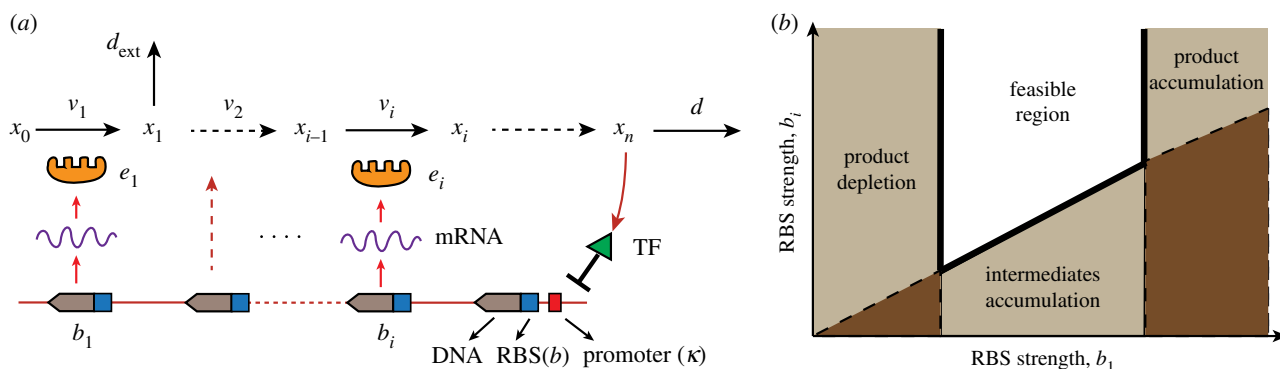
**Figure 5.** Control system structure of the ATP energy metabolism network of figure 3.  $k_a$  is the protein synthesis rate constant.  $k_0$  and  $k_b$  are the zero- and first-order protein degradation rate constants.  $k_d$  is the perturbation. (Online version in colour.)

control gene expression dynamically. Assuming all the enzymes of an unbranched pathway to be encoded in an operon under the control of a single promoter (figure 6a), the expression of catalytic enzymes can be expressed as

$$\dot{e}_i = b_i \cdot (\kappa^0 + \kappa^1 \cdot \sigma(x_n)) - k_i^d \cdot e_i. \quad (4.1)$$

Here  $i = 1, 2, \dots, n$ . The functional  $\sigma(\cdot)$  represents the metabolite–TF and TF–promoter feedback regulation.  $\kappa^0$  and  $\kappa^1$  denote the promoter tightness and strength, and therewith describe the regulatory effect of TF on gene transcription. The translation rate of enzymes can be modified by choosing appropriate RBS strength  $b_i$ , while the protein degradation rate can be controlled by adding a degradation tag to the gene sequence or altering synonymous codons [138]. Recently, two engineered sensor-regulator systems were built to improve fatty acid-based biofuel production, where several heterologous genes and a flux-controlling enzyme acetyl-CoA carboxylase were controlled dynamically by key metabolites, acetyl-CoA and malonyl-CoA, respectively [37,139]. Additionally, a genetically encoded synthetic malonyl-CoA switch was developed recently [140]. In such engineered feedback sensor-regulator systems, a transcription factor is often identified as a natural sensor that senses the biosynthetic intermediate. The naturally occurring cognate regulators or controllers (e.g. regulatory DNA elements) may need to be re-engineered for use with natural sensors.

Classical MCA has not yet been used much for the design of circuits in synthetic biology, but MCA and HCA can be used to analyse the effects on productivity of modifying parameters (e.g. promoter strengths or protein degradation tag) of both natural and synthetic gene circuits at the DNA, mRNA and protein levels. Hence, MCA/HCA can also assist in the design of more robust metabolic pathways because of their strength in quantifying network robustness as discussed in previous sections. HCA applies to any dynamic network, inclusive of those involved in transcriptional and translational regulation [117,141]. By viewing the level of mRNA as a resultant from a synthesis and a degradation process, HCA (or even MCA) predicts how promoter strength or degradation activity controls the level of the mRNA (e.g. [142]). This has been extended to the intricacies of transcriptional regulation in mammalian systems in the context of epigenetics and chromatin modification [143], although more work is needed here. Mathematically, the dynamics of a synthetic gene-metabolic pathway can be expressed by combining the expression of (2.1) and (4.1), and we can further break down the dynamics of the protein concentrations (4.1) into transcription (mRNA dynamics) and translation. As a result, this HCA can be used to track how



**Figure 6.** Important tunable ‘design parameters’ of synthetic gene circuits and design constraints of the RBS strength for a metabolic pathway with a branch consuming an intermediate. (a) The promoter characteristic ( $\kappa^0, \kappa^1$ ) and RBS strengths ( $b_i$ ) can be re-engineered to modulate gene transcription and translation rates. (b) Design constraints with respect to RBS ratio  $b_i/b_1$ ; the feasible RBS region prevents the accumulation of product and intermediates as well as avoids product depletion due to the weak control of flux through the branch. Adapted from [39]. (Online version in colour.)

the genetic parameters (such as the promoter strength  $\kappa^1$ ) control the mRNA level or protein level  $E$  (or even flux) as

$$H_{\kappa^1}^{[\text{mRNA}]} = \frac{\delta \ln[\text{mRNA}]}{\delta \ln \kappa^1} \quad \text{or} \quad H_{\kappa^1}^E = \frac{\delta \ln E}{\delta \ln \kappa^1}. \quad (4.2)$$

This is similar to the hierarchical control of enzyme level discussed for the ATP energy metabolism example (§3.2) and the concentration control coefficients used in MCA, although at a different biological network level.

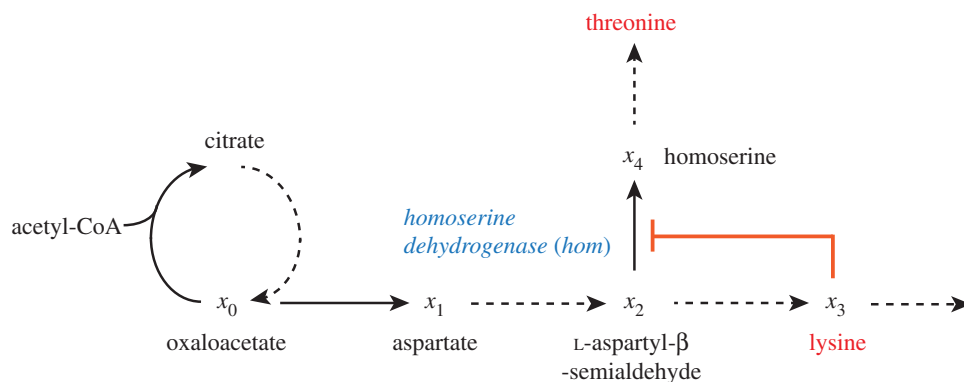
Before implementing an engineered gene circuit, the design constraints with respect to design parameters need to be identified. Oyarzun & Stan [39] investigated the trade-offs and constraints in designing the gene-expression regulation circuits in unbranched pathways under the single operon assumption as expressed in (4.1). Constraints to tunable design parameters, such as the promoter’s dynamic characteristics and RBS strengths (figure 6a), must be satisfied to guarantee the existence of a stable steady state and to prevent the continued accumulation of intermediates (as shown in figure 6b). The local stability of the metabolic-regulatory system (i.e. combining (2.1) and (4.1)) under a perturbation, such as a change in cellular demand, was studied by applying a systems engineering approach in terms of the design elements in the gene circuits. In addition, it was concluded that the use of promoters with a broad dynamic range (e.g. increased tightness) can enlarge the feasible design region, whereas leakier promoters or higher substrate concentrations tighten the constraints. This is similar to the design of a mechanical or electronic control system. The stability of the closed-loop control system together with the actuator or output and state variable constraints need to be considered during the controller design process.

The design constraints of a synthetic metabolism at a pathway level were also studied [144] by combining FBA and Markov chain Monte Carlo sampling in a ‘global’ metabolic (genotype) network space. Specifically, this study investigated the quantitative relationships between different properties in a network, i.e. the number of alternative carbon sources it can use, the amount of biomass it can synthesize, the number of active reactions and the amount of waste the network produces, as well as how these metabolic network properties influence biosynthetic flux.

## 4.2. Synthetic allostery and protein–protein association

Apart from synthetic gene circuits for gene-expression regulation, post-translational control of the native allostery in metabolic regulation can also be re-engineered and optimized to provide a faster way for the control of metabolic pathways [145]. The allosteric network structure can be reconstructed to optimize a metabolic control, for example, eliminating allosteric enzyme inhibition so as to maximize product yield, or integrating allosteric inhibition into enzymes so as to increase product purity by eliminating by-product synthesis pathways [146]. The mechanism of allostery (i.e. structural and dynamic information related to the allosteric sites on the enzymes) can also be modified and fine-tuned to provide a more accurate and moderate regulation. Statistical coupling analysis and molecular dynamic simulation of proteins have been combined to re-engineer the allosteric regulation of aspartokinase, in *E. coli* and *C. glutamicum*, so that a reduced allosteric inhibition of the production of lysine and/or an enhanced inhibition of the formation of the by-product threonine was achieved [147] as shown in figure 7.

Another attractive strategy for balancing pathway fluxes is to design scaffold protein [148] which then co-localizes cascade enzymes which results in synthetic protein–protein complex formation and substrate channelling. Substrate channelling has important potential applications in metabolic engineering, such as decreasing the effective free concentrations of unstable intermediates [149], preventing the accumulation of toxic metabolites, regulating fluxes by modulating enzyme–enzyme association and increasing sensitivity to regulatory signals [150]. Some examples have been reported, such as three non-endogenous mevalonate biosynthesis enzymes being scaffolded in *E. coli* using eukaryote protein–protein interaction domains [148]. Immobilized three-enzyme complexes (named synthetic metabolon), i.e. triosephosphate isomerase, aldolase and fructose 1,6-bisphosphatase in the glycolysis and gluconeogenesis pathways, have been constructed, which not only decreased protein purification labour and cost but also accelerated reaction rates by an order of magnitude when compared with non-complexed enzymes [151]. The flux control for a pathway with metabolic channelling at steady state was theoretically quantified under an MCA framework [152]. Quantitative analysis of substrate channelling and scaffold proteins on metabolic network dynamics, especially in a systems engineering context, and corresponding synthetic design constraints remain open research topics.



**Figure 7.** Illustration of a simplified metabolic pathway and control of the by-product (threonine) synthesis for increased lysine production in bacteria. A non-natural allosteric regulation of homoserine dehydrogenase in response to lysine was constructed to reduce the flux to by-product threonine. (Online version in colour.)

### 4.3. Modularity and retroactivity

A complex biochemical network is often decomposed into simpler components or into modules with common functions in order to facilitate the analysis and prediction of the overall system dynamic behaviour. On the other hand, the synthetic modular ‘circuits’ composed of genes and proteins are also required to be placed into living cells (through the process of transformation or transfection). A fundamental systems engineering issue then arises when interconnecting several subsystems that is: whether and how the transmitting of a signal to a ‘downstream’ subsystem affects the dynamic behaviour of the ‘upstream’ subsystem. Recent studies on signalling and transcriptional network modules [38,153], such as the covalent modification cascades (e.g. MAPK cascades) and the machinery for protein synthesis (TF–protein interactions), reveal that modularity does not hold in general. Both an impedance-like effect that is well known in electrical systems and called ‘retroactivity’ [38], and unintended ‘crosstalk’ [154,155] have been identified and quantified in biochemical networks. They are at the very basis of the complexity of most biological networks that produces their circular causality, making an effect the cause of its cause [156]. Several approaches have been proposed to attenuate the effect of retroactivity or crosstalk, for example, inserting insulators between modules, layering methods, scaffolding [148], introducing negative feedback mechanisms (inspired by the design of amplifiers in electronics) through transcriptional activation or through phosphorylation of a protein that is in abundance [38], or using a classical loop-shaping control design [157] as an alternative to inserting insulators.

In a metabolic network, the metabolic supply and demand analysis (as discussed in §2.3) studies how the steady-state properties of a supply process can be controlled by the demand process that consumes the product. This is related to the retroactivity concept and analysis, and the latter concerns not only the steady-state behaviour but also transient situations. Similarly, modular MCA (or HCA) [74,128] may be used to quantify the retroactivity at the steady state that requires the measuring of the ‘overall’ control exerted by the processes catalysed by each module.

The retroactivity in gene transcriptional and signal-transduction networks has been formulated into a general control theoretic, i.e. disturbance rejection, problem [38,158]. It will be exciting to investigate the retroactivity problem of a metabolic network subject to both gene-expression and allosteric regulatory modules in terms of a generic control-

theoretic problem and to study how it is related to MCA (and HCA) and supply–demand analysis.

### 4.4. Noise propagation and stochasticity

Gene expression is known to be noisy as transcription and translation may depend on infrequent events involving small numbers of molecules, as well as environmental fluctuations and genetic mutation [159]. There are many studies on the noise and stochastic effects in individual genes and gene regulatory circuits [160,161]. The retroactivity effects discussed in the previous subsection can also be practically inferred from the stochastic noise effect, because the ‘noise’ can be an information-rich ‘input’ signal designed to study the changes in the systems dynamic responses [162]. Stochastic effects in metabolic pathways were initially studied with respect to fluctuations in substrates and upstream metabolites [163]. Noise propagation from enzyme expression is typically neglected because of the assumptions on high metabolite counts and because stochasticity in enzyme expression occurs at much longer time scales. The above assumptions however may not hold for engineered metabolic pathways that can operate at low metabolite counts. Recently, simulation studies on noise propagation from enzyme expression to simple metabolic pathways [164,165] have been provided. How the extrinsic fluctuations (at the kinetic level) can affect the steady-state control of pathway fluxes and concentrations was investigated based on the stochastic generalizations of standard MCA [166,167]. Based on stochastic simulations, Oyarzun *et al.* [168] quantified the noise propagation in a simple synthetic metabolic pathway under transcriptional repression by the product and its dependency on the design parameters, such as promoter strength, promoter sensitivity and repression strength. It was concluded that attenuation of metabolic noise is a trade-off between the promoter strength and repression strength.

## 5. Conclusion and perspectives

Metabolic engineering has been applied to the analysis, design and optimization of metabolic pathways for more than three decades. Successful applications of synthetic biology on metabolic processes have been reported in the past decade, while the use of control engineering approaches to optimize, analyse and assist the synthetic biology design of metabolic networks has only begun to attract the attention in

the last few years and may still be in its infancy. This review discussed important metabolic engineering approaches and synthetic biology-related design problems for metabolic-regulatory networks, and their connections with systems and control engineering.

Control theory can advance classical metabolic engineering techniques in pathway optimization, control and regulation analysis. To optimize the flux distribution and the production of a pathway while taking the network dynamics into account, the dynamic FBA can be formulated into an optimal control problem. As long as the objective function is linear or quadratic and the metabolic network is subject to only dynamic constraints, analytical solutions of such DFBA can be obtained from linear optimal control theory (e.g. the LQR approach mentioned in §3.1). When the objective function or system dynamics are nonlinear and the network is subject to certain constraints, such as in the dynamic enzyme activation problem, advanced nonlinear optimal control strategies need to be employed to solve such nonlinear optimization problems. One important future piece of work is to accommodate the stochastic effects and uncertainties in an FBA model, such as the uncertainties in the measured biomass composition or the flux capacities [169]. Robust control and optimal control theories are powerful tools that may be used to tackle this problem.

Apart from the optimization of flux distribution, the steady-state control of a network parameter (e.g. enzyme concentration or activity) on pathway fluxes or metabolite concentrations, or vice versa regulatory effects such as the inhibition of enzyme activities via allosteric regulation, are traditionally quantified by MCA and regulation analysis, respectively. Both control theory and HCA can generalize the MCA-based analysis to dynamic situations and take account of different regulation mechanisms at different hierarchical levels, such as the enzymatic and gene-expression regulation. These can be linked with the classical proportional and integral control mechanisms of control engineering. As a result, the enzymatic control, regulation and the cellular robust adaptation to environmental conditions can all be studied within the well-established control-theoretic framework both for steady state and for transient states.

Successful synthetic biology applications based on metabolic-regulatory networks have been reported in the past 10 years; the use of control engineering techniques to assist the synthetic design is more recent and less complete. It includes

(i) the use of feedback control in the design and realization of synthetic gene-metabolic systems, including the implementation of biological sensors, controllers and actuators, (ii) the clarification of different metabolic or genetic regulatory mechanisms in terms of widely used control mechanisms [35,77], (iii) attenuation of the retroactivity, crosstalk and noise effects when interconnecting synthetic modules by introducing feedback loops around modules, and (iv) identification of the design constraints and of the existence of local stability in terms of design parameters such as promoter and ribosome characteristics.

It can be expected that systems and control engineering will now serve to advance existing metabolic engineering and synthetic biology techniques to improve metabolic product and network robustness. Challenges will be due to the nonlinear and complex nature of the metabolic networks and to the complicated regulatory mechanisms at different hierarchical levels. Important future work includes: (i) design and implementation of various new control strategies (i.e. a mixture of proportional, integral, derivative or nonlinear controllers) in a complex bio-synthetic pathway with regulation at hierarchical (metabolic, gene expression and signal transduction) levels, (ii) design and implementation of bio-controllers that can accommodate network parametric uncertainties, which can be achieved either through quantifying the uncertainties in the design process by the use of robust control theory [170], or by engineering and inserting a non-natural bio-integrator into the metabolic-regulatory system, since an integral feedback control is known to achieve perfect robustness and adaptation, and (iii) rigorous analysis of the design constraints and network stability when adding a synthetic genetic or metabolic module to the overall network.

Much remains to be discovered and then engineered.

**Authors' contributions.** F.H. and H.V.W. generated the ideas for the manuscript, helped to draft and edit it. F.H. performed literature searches. E.M. contributed to the section on FBA. H.V.W. helped with the final editing done by F.H.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was supported various systems biology grants, including Synpol: EU-FP7 (KBBE.2012.3.4-02 no. 311815), Corbel: EU-H2020 (NFRADEV-4-2014-2015 no. 654248), EpiPredict: EU-H2020 MSCA-ITN-2014-ETN: Marie Skłodowska-Curie Innovative Training Networks (ITN-ETN) no. 642691, BBSRC China: BB/J020060/1.

**Acknowledgements.** We thank anonymous reviewers for pointing out some additional issues and work to be referred to.

## References

- Bailey JE. 1991 Toward a science of metabolic engineering. *Science* **252**, 1668–1675. (doi:10.1126/science.2047876)
- Thiele I *et al.* 2013 A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **31**, 419–425. (doi:10.1038/nbt.2488)
- Stephanopoulos G. 1999 Metabolic fluxes and metabolic engineering. *Metab. Eng.* **1**, 1–11. (doi:10.1006/mben.1998.0101)
- Kacser H, Burns JA, Fell DA. 1995 The control of flux. *Biochem. Soc. Trans.* **23**, 341–366. (doi:10.1042/bst0230341)
- Fell DA. 1997 *Understanding the control of metabolism*. London, UK: Portland Press.
- Kahn D, Westerhoff HV. 1991 Control-theory of regulatory cascades. *J. Theor. Biol.* **153**, 255–285. (doi:10.1016/S0022-5193(05)80426-6)
- Orth JD, Thiele I, Palsson BO. 2010 What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248. (doi:10.1038/nbt.1614)
- Monk J, Nogales J, Palsson BO. 2014 Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* **32**, 447–452. (doi:10.1038/nbt.2870)
- McCloskey D, Palsson BO, Feist AM. 2013 Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol. Syst. Biol.* **9**, 661. (doi:10.1038/msb.2013.18)
- Alper H, Jin YS, Moxley JF, Stephanopoulos G. 2005 Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.* **7**, 155–164. (doi:10.1016/j.mben.2004.12.003)
- McAnulty MJ, Yen JY, Freedman BG, Senger RS. 2012 Genome-scale modeling using flux ratio constraints to enable metabolic engineering of

- clostridial metabolism *in silico*. *BMC Syst. Biol.* **6**, 42. (doi:10.1186/1752-0509-6-42)
12. Blazeck J, Alper H. 2010 Systems metabolic engineering: genome-scale models and beyond. *Biotechnol. J.* **5**, 647–659. (doi:10.1002/biot.200900247)
  13. Martin CH, Nielsen DR, Solomon KV, Prather KLJ. 2009 Synthetic metabolism: engineering biology at the protein and pathway scales. *Chem. Biol.* **16**, 277–286. (doi:10.1016/j.chembiol.2009.01.010)
  14. Andrianantoandro E, Basu S, Karig DK, Weiss R. 2006 Synthetic biology: new engineering rules for an emerging discipline. *Mol. Syst. Biol.* **2**, 28. (doi:10.1038/Msb4100073)
  15. Cheng AA, Lu TK. 2012 Synthetic biology: an emerging engineering discipline. *Annu. Rev. Biomed. Eng.* **14**, 155–178. (doi:10.1146/annurev-bioeng-071811-150118)
  16. Boyle PM, Silver PA. 2009 Harnessing nature's toolbox: regulatory elements for synthetic biology. *J. R. Soc. Interface* **6**(Suppl. 4), S535–S546. (doi:10.1098/rsif.2008.0521.focus)
  17. Slusarczyk AL, Lin A, Weiss R. 2012 Foundations for the design and implementation of synthetic genetic circuits. *Nat. Rev. Genet.* **13**, 406–420. (doi:10.1038/nrg3227)
  18. Khalil AS, Collins JJ. 2010 Synthetic biology: applications come of age. *Nat. Rev. Genet.* **11**, 367–379. (doi:10.1038/Nrg2775)
  19. Purnick PE, Weiss R. 2009 The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* **10**, 410–422. (doi:10.1038/nrm2698)
  20. Chen BS, Wu CH. 2009 A systematic design method for robust synthetic biology to satisfy design specifications. *BMC Syst. Biol.* **3**, 66. (doi:10.1186/1752-0509-3-66)
  21. Wang YH, Wei KY, Smolke CD. 2013 Synthetic biology: advancing the design of diverse genetic systems. *Annu. Rev. Chem. Biomol.* **4**, 69–102. (doi:10.1146/annurev-chembioeng-061312-103351)
  22. Barrett CL, Kim TY, Kim HU, Palsson BO, Lee SY. 2006 Systems biology as a foundation for genome-scale synthetic biology. *Curr. Opin. Biotechnol.* **17**, 488–492. (doi:10.1016/j.copbio.2006.08.001)
  23. Farmer WR, Liao JC. 2000 Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nat. Biotechnol.* **18**, 533–537. (doi:10.1038/75398)
  24. Fung E, Wong WW, Suen JK, Bulter T, Lee SG, Liao JC. 2005 A synthetic gene-metabolic oscillator. *Nature* **435**, 118–122. (doi:10.1038/nature03508)
  25. Stephanopoulos G. 2012 Synthetic biology and metabolic engineering. *ACS Synth. Biol.* **1**, 514–525. (doi:10.1021/Sb300094q)
  26. Keasling JD. 2012 Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* **14**, 189–195. (doi:10.1016/j.ymben.2012.01.004)
  27. Yadav VG, De Mey M, Lim CG, Ajikumar PK, Stephanopoulos G. 2012 The future of metabolic engineering and synthetic biology: towards a systematic practice. *Metab. Eng.* **14**, 233–241. (doi:10.1016/j.ymben.2012.02.001)
  28. Boyle PM, Silver PA. 2012 Parts plus pipes: synthetic biology approaches to metabolic engineering. *Metab. Eng.* **14**, 223–232. (doi:10.1016/j.ymben.2011.10.003)
  29. Rabinovitch-Deere CA, Oliver JWK, Rodriguez GM, Atsumi S. 2013 Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* **113**, 4611–4632. (doi:10.1021/cr300361t)
  30. Shen-Orr SS, Milo R, Mangan S, Alon U. 2002 Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* **31**, 64–68. (doi:10.1038/ng881)
  31. Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. 2009 Reconstruction of biochemical networks in microorganisms. *Nat. Rev. Microbiol.* **7**, 129–143. (doi:10.1038/Nrmicro1949)
  32. Bier M, Bakker BM, Westerhoff HV. 2000 How yeast cells synchronize their glycolytic oscillations: a perturbation analytic treatment. *Biophys. J.* **78**, 1087–1093. (doi:10.1016/S0006-3495(00)76667-7)
  33. Cury JER, Baldissera FL. 2013 Systems biology, synthetic biology and control theory: a promising golden braid. *Annu. Rev. Control* **37**, 57–67. (doi:10.1016/j.arcontrol.2013.03.006)
  34. Prescott TP, Papachristodoulou A. 2014 Synthetic biology: a control engineering perspective. In *European Control Conf., Strasbourg, France, 24–27 June 2014*, pp. 1182–1186. (doi:10.1109/ECC.2014.6862638)
  35. Cloutier M, Wellstead P. 2010 The control systems structures of energy metabolism. *J. R. Soc. Interface* **7**, 651–665. (doi:10.1098/rsif.2009.0371)
  36. Michener JK, Thodey K, Liang JC, Smolke CD. 2012 Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metab. Eng.* **14**, 212–222. (doi:10.1016/j.ymben.2011.09.004)
  37. Zhang FZ, Carothers JM, Keasling JD. 2012 Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **30**, 354–359. (doi:10.1038/nbt.2149)
  38. Del Vecchio D, Ninfa AJ, Sontag ED. 2008 Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* **4**, 161. (doi:10.1038/msb4100204)
  39. Oyarzun DA, Stan GB. 2012 Synthetic gene circuits for metabolic control: design trade-offs and constraints. *J. R. Soc. Interface* **10**, 20120671. (doi:10.1098/rsif.2012.0671)
  40. Hooshangi S, Thiberge S, Weiss R. 2005 Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl Acad. Sci. USA* **102**, 3581–3586. (doi:10.1073/pnas.0408507102)
  41. Bennett MR, Pang WL, Ostroff NA, Baumgartner BL, Nayak S, Tsimring LS, Hasty J. 2008 Metabolic gene regulation in a dynamically changing environment. *Nature* **454**, 1119–1122. (doi:10.1038/nature07211)
  42. Moreno-Sanchez R, Saavedra E, Rodriguez-Enriquez S, Olin-Sandoval V. 2008 Metabolic control analysis: a tool for designing strategies to manipulate metabolic pathways. *J. Biomed. Biotechnol.* **2008**, 597913. (doi:10.1155/2008/597913)
  43. Raman K, Chandra N. 2009 Flux balance analysis of biological systems: applications and challenges. *Brief. Bioinform.* **10**, 435–449. (doi:10.1093/bib/bbp011)
  44. Lourenco A, Carneiro S, Pinto JP, Rocha M, Ferreira EC, Rocha I. 2011 A study of the short and long-term regulation of *E. coli* metabolic pathways. *J. Integr. Bioinform.* **8**, 183. (doi:10.2390/biecoll-jib-2011-183)
  45. van Heeswijk WC, Westerhoff HV, Booger FC. 2013 Nitrogen assimilation in *Escherichia coli*: putting molecular data into a systems perspective. *Microbiol. Mol. Biol. Rev.* **77**, 628–695. (doi:10.1128/Mmbr.00025-13)
  46. Rossell S, van der Weijden CC, Lindenberg A, van Tuijl A, Francke C, Bakker BM, Westerhoff HV. 2006 Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **103**, 2166–2171. (doi:10.1073/pnas.0509831103)
  47. Chubukov V, Gerosa L, Kochanowski K, Sauer U. 2014 Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **12**, 327–340. (doi:10.1038/nrmicro3238)
  48. Alon U. 2007 Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450–461. (doi:10.1038/nrg2102)
  49. Yeger-Lotem E, Sattath S, Kashtan N, Itzkovitz S, Milo R, Pinter RY, Alon U, Margalit H. 2004 Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction. *Proc. Natl Acad. Sci. USA* **101**, 5934–5939. (doi:10.1073/pnas.0306752101)
  50. Rossell S, van der Weijden CC, Kruckeberg AL, Bakker BM, Westerhoff HV. 2005 Hierarchical and metabolic regulation of glucose influx in starved *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **5**, 611–619. (doi:10.1016/j.femsyr.2004.11.003)
  51. Daran-Lapujade P *et al.* 2007 The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *Proc. Natl Acad. Sci. USA* **104**, 15 753–15 758. (doi:10.1073/pnas.0707476104)
  52. Herrgard MJ *et al.* 2008 A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* **26**, 1155–1160. (doi:10.1038/nbt1492)
  53. Burgard AP, Pharkya P, Maranas CD. 2003 OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657. (doi:10.1002/bit.10803)
  54. Chowdhury A, Zomorodi AR, Maranas CD. 2014 k-OptForce: integrating kinetics with flux balance analysis for strain design. *PLoS Comput. Biol.* **10**, e1003487. (doi:10.1371/journal.pcbi.1003487)
  55. Kim J, Reed JL, Maravelias CT. 2011 Large-scale bilevel strain design approaches and mixed-integer programming solution techniques. *PLoS ONE* **6**, e0024162. (doi:10.1371/journal.pone.0024162)
  56. Burgard AP, Vaidyaraman S, Maranas CD. 2001 Minimal reaction sets for *Escherichia coli*



- metabolism under different growth requirements and uptake environments. *Biotechnol. Progr.* **17**, 791–797. (doi:10.1021/Bp0100880)
57. Mahadevan R, Schilling CH. 2003 The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **5**, 264–276. (doi:10.1016/j.ymben.2003.09.002)
  58. Covert MW, Schilling CH, Palsson B. 2001 Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* **213**, 73–88. (doi:10.1006/jtbi.2001.2405)
  59. Covert MW, Palsson BO. 2002 Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* **277**, 28 058–28 064. (doi:10.1074/jbc.M201691200)
  60. Gianchandani EP, Chavali BK, Papin JA. 2010 The application of flux balance analysis in systems biology. *Wires Syst. Biol. Med.* **2**, 372–382. (doi:10.1002/Wsbm.60)
  61. Gianchandani EP, Papin JA, Price ND, Joyce AR, Palsson BO. 2006 Matrix formalism to describe functional states of transcriptional regulatory systems. *PLoS Comput. Biol.* **2**, e101. (doi:10.1371/journal.pcbi.0020101)
  62. Schuetz R, Kuepfer L, Sauer U. 2007 Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119. (doi:10.1038/msb4100162)
  63. Schuster S, Pfeiffer T, Fell DA. 2008 Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.* **252**, 497–504. (doi:10.1016/j.jtbi.2007.12.008)
  64. Simeonidis E, Murabito E, Smallbone K, Westerhoff HV. 2010 Why does yeast ferment? A flux balance analysis study. *Biochem. Soc. Trans.* **38**, 1225–1229. (doi:10.1042/Bst0381225)
  65. Shlomi T, Cabili MN, Herrgard MJ, Palsson BO, Ruppin E. 2008 Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **26**, 1003–1010. (doi:10.1038/nbt.1487)
  66. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, Mendes P, Swainston N. 2012 Improving metabolic flux predictions using absolute gene expression data. *BMC Syst. Biol.* **6**, 73. (doi:10.1186/1752-0509-6-73)
  67. Westerhoff HV. 2008 Signalling control strength. *J. Theor. Biol.* **252**, 555–567. (doi:10.1016/j.jtbi.2007.11.035)
  68. Mahadevan R, Edwards JS, Doyle FJ. 2002 Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* **83**, 1331–1340. (doi:10.1016/S0006-3495(02)73903-9)
  69. Grafahrend-Belau E, Junker A, Eschenroder A, Muller J, Schreiber F, Junker BH. 2013 Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. *Plant Physiol.* **163**, 637–647. (doi:10.1104/pp.113.224006)
  70. Waldherr S, Oyarzun DA, Bockmayr A. 2015 Dynamic optimization of metabolic networks coupled with gene expression. *J. Theor. Biol.* **365**, 469–485. (doi:10.1016/j.jtbi.2014.10.035)
  71. Rugen M, Bockmayr A, Steuer R. 2015 Elucidating temporal resource allocation and diurnal dynamics in phototrophic metabolism using conditional FBA. *Sci. Rep.-UK* **5**, 15247. (doi:10.1038/srep15247)
  72. Westerhoff HV, Dam K. 1987 *Thermodynamics and control of biological free-energy transduction*. Amsterdam, The Netherlands: Elsevier.
  73. Kholodenko BN, Molenaar D, Schuster S, Heinrich R, Westerhoff HV. 1995 Defining control coefficients in nonideal metabolic pathways. *Biophys. Chem.* **56**, 215–226. (doi:10.1016/0301-4622(95)00039-Z)
  74. Schuster S, Kahn D, Westerhoff HV. 1993 Modular analysis of the control of complex metabolic pathways. *Biophys. Chem.* **48**, 1–17. (doi:10.1016/0301-4622(93)80037-J)
  75. Hofmeyr JHS, Cornish-Bowden A. 1991 Quantitative assessment of regulation in metabolic systems. *Eur. J. Biochem.* **200**, 223–236. (doi:10.1111/j.1432-1033.1991.tb21071.x)
  76. Hofmeyr JHS, Cornish-Bowden A. 2000 Regulating the cellular economy of supply and demand. *FEBS Lett.* **476**, 47–51. (doi:10.1016/S0014-5793(00) 01668-9)
  77. He F, Fromion V, Westerhoff HV. 2013 (Im)Perfect robustness and adaptation of metabolic networks subject to metabolic and gene-expression regulation: marrying control engineering with metabolic control analysis. *BMC Syst. Biol.* **7**, 131. (doi:10.1186/1752-0509-7-131)
  78. Miskovic L, Hatzimanikatis V. 2011 Modeling of uncertainties in biochemical reactions. *Biotechnol. Bioeng.* **108**, 413–423. (doi:10.1002/bit.22932)
  79. Saa P, Nielsen LK. 2015 A general framework for thermodynamically consistent parameterization and efficient sampling of enzymatic reactions. *PLoS Comput. Biol.* **11**, e1004195. (doi:10.1371/journal.pcbi.1004195)
  80. Chakrabarti A, Miskovic L, Soh KC, Hatzimanikatis V. 2013 Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints. *Biotechnol. J.* **8**, 1043–1105. (doi:10.1002/biot.201300091)
  81. Oliveira AP, Ludwig C, Zampieri M, Weisser H, Aebersold R, Sauer U. 2015 Dynamic phosphoproteomics reveals TORC1-dependent regulation of yeast nucleotide and amino acid biosynthesis. *Sci. Signal* **8**, rs4. (doi:10.1126/scisignal.2005768)
  82. Kochanowski K, Sauer U, Noor E. 2015 Posttranslational regulation of microbial metabolism. *Curr. Opin. Microbiol.* **27**, 10–17. (doi:10.1016/j.mib.2015.05.007)
  83. Kerkhoven EJ, Lahtvee PJ, Nielsen J. 2015 Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res.* **15**, 1–13. (doi:10.1111/1567-1364.12199)
  84. Westerhoff HV *et al.* 2009 Systems biology towards life *in silico*: mathematics of the control of living cells. *J. Math. Biol.* **58**, 7–34. (doi:10.1007/s00285-008-0160-8)
  85. Sauro HM. 1990 Regulatory responses and control analysis—assessment of the relative importance of internal effectors. In *Control of metabolic processes* (ed. A Cornish-Bowden), pp. 225–230. New York, NY: Plenum Press.
  86. ter Kuile BH, Westerhoff HV. 2001 Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett.* **500**, 169–171. (doi:10.1016/S0014-5793(01) 02613-8)
  87. Even S, Lindley ND, Coccagn-Bousquet M. 2003 Transcriptional, translational and metabolic regulation of glycolysis in *Lactococcus lactis* subsp. cremoris MG 1363 grown in continuous acidic cultures. *Microbiology* **149**, 1935–1944. (doi:10.1099/mic.0.26146-0)
  88. Bruggeman FJ, Boogerd FC, Westerhoff HV. 2005 The multifarious short-term regulation of ammonium assimilation of *Escherichia coli*: dissection using an *in silico* replica. *FEBS J.* **272**, 1965–1985. (doi:10.1111/j.1742-4658.2005. 04626.x)
  89. Bruggeman FJ, Westerhoff HV. 2007 The nature of systems biology. *Trends Microbiol.* **15**, 45–50. (doi:10.1016/j.tim.2006.11.003)
  90. Weinert BT, Iesmantavicius V, Wagner SA, Scholz C, Gummesson B, Beli P, Nystrom T, Choudhary C. 2013 Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* **51**, 265–272. (doi:10.1016/j.molcel.2013.06.003)
  91. Moellering RE, Cravatt BF. 2013 Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. *Science* **341**, 549–553. (doi:10.1126/science.1238327)
  92. Bruggeman FJ, de Haan J, Hardin H, Bouwman J, Rossell S, van Eunen K, Bakker BM, Westerhoff HV. 2006 Time-dependent hierarchical regulation analysis: deciphering cellular adaptation. *IEE Proc. Syst. Biol.* **153**, 318–322. (doi:10.1049/ip-syb:20060027)
  93. van Eunen K, Bouwman J, Lindenberg A, Westerhoff HV, Bakker BM. 2009 Time-dependent regulation analysis dissects shifts between metabolic and gene-expression regulation during nitrogen starvation in baker's yeast. *FEBS J.* **276**, 5521–5536. (doi:10.1111/j.1742-4658.2009. 07235.x)
  94. Westerhoff HV, Verma M, Nardelli M, Adamczyk M, van Eunen K, Simeonidis E, Bakker BM. 2010 Systems biochemistry in practice: experimenting with modelling and understanding, with regulation and control. *Biochem. Soc. Trans.* **38**, 1189–1196. (doi:10.1042/Bst0381189)
  95. Albe KR, Wright BE. 1992 Systems analysis of the tricarboxylic acid cycle in *Dictyostelium discoideum*. II. Control analysis. *J. Biol. Chem.* **267**, 3106–3114.
  96. Araujo WL, Nunes-Nesi A, Nikoloski Z, Sweetlove LJ, Fernie AR. 2012 Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. *Plant Cell Environ.* **35**, 1–21. (doi:10.1111/j.1365-3040.2011.02332.x)
  97. Behre J, Wilhelm T, von Kamp A, Ruppin E, Schuster S. 2008 Structural robustness of metabolic networks with respect to multiple knockouts. *J. Theor. Biol.* **252**, 433–441. (doi:10.1016/j.jtbi. 2007.09.043)
  98. Kwon YK, Cho KH. 2008 Quantitative analysis of robustness and fragility in biological networks

- based on feedback dynamics. *Bioinformatics* **24**, 987–994. (doi:10.1093/bioinformatics/btn060)
99. Stelling J, Sauer U, Szallasi Z, Doyle FJ, Doyle J. 2004 Robustness of cellular functions. *Cell* **118**, 675–685. (doi:10.1016/j.cell.2004.09.008)
  100. Quinton-Tulloch MJ, Bruggeman FJ, Snoep JL, Westerhoff HV. 2013 Trade-off of dynamic fragility but not of robustness in metabolic pathways *in silico*. *FEBS J.* **280**, 160–173. (doi:10.1111/febs.12057)
  101. Cascante M, Boros LG, Comin-Anduix B, de Atauri P, Centelles JJ, Lee PWN. 2002 Metabolic control analysis in drug discovery and disease. *Nat. Biotechnol.* **20**, 243–249. (doi:10.1038/Nbt0302-243)
  102. Bakker BM, Westerhoff HV, Opperdoes FR, Michels PA. 2000 Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs. *Mol. Biochem. Parasitol.* **106**, 1–10. (doi:10.1016/S0166-6851(99)00197-8)
  103. Murabito E, Smallbone K, Swinton J, Westerhoff HV, Steuer R. 2011 A probabilistic approach to identify putative drug targets in biochemical networks. *J. R. Soc. Interface* **8**, 880–895. (doi:10.1098/rsif.2010.0540)
  104. Snoep JL, van der Weijden CC, Andersen HW, Westerhoff HV, Jensen PR. 2002 DNA supercoiling in *Escherichia coli* is under tight and subtle homeostatic control, involving gene-expression and metabolic regulation of both topoisomerase I and DNA gyrase. *Eur. J. Biochem.* **269**, 1662–1669. (doi:10.1046/j.1432-1327.2002.02803.x)
  105. Acerenza L, Sauro HM, Kacser H. 1989 Control analysis of time-dependent metabolic systems. *J. Theor. Biol.* **137**, 423–444. (doi:10.1016/S0022-5193(89)80038-4)
  106. Kholodenko BN, Demin OV, Westerhoff HV. 1997 Control analysis of periodic phenomena in biological systems. *J. Phys. Chem. B* **101**, 2070–2081. (doi:10.1021/Jp962336u)
  107. Demin OV, Westerhoff HV, Kholodenko BN. 1999 Control analysis of stationary forced oscillations. *J. Phys. Chem. B* **103**, 10 695–10 710. (doi:10.1021/jp991597b)
  108. Reijenga KA, Westerhoff HV, Kholodenko BN, Snoep JL. 2002 Control analysis for autonomously oscillating biochemical networks. *Biophys. J.* **82**, 99–108. (doi:10.1016/S0006-3495(02)75377-0)
  109. Gustavsson AK, van Niekerk DD, Adiels CB, Kooi B, Goksor M, Snoep JL. 2014 Allosteric regulation of phosphofructokinase controls the emergence of glycolytic oscillations in isolated yeast cells. *FEBS J.* **281**, 2784–2793. (doi:10.1111/febs.12820)
  110. Conradie R, Bruggeman FJ, Ciliberto A, Csikasz-Nagy A, Novak B, Westerhoff HV, Snoep JL. 2010 Restriction point control of the mammalian cell cycle via the cyclin E/Cdk2:p27 complex. *FEBS J.* **277**, 357–367. (doi:10.1111/j.1742-4658.2009.07473.x)
  111. Hornberg JJ, Bruggeman FJ, Binder B, Geest CR, de Vaate AJMB, Lankelma J, Heinrich R, Westerhoff HV. 2005 Principles behind the multifarious control of signal transduction—ERK phosphorylation and kinase/phosphatase control. *FEBS J.* **272**, 244–258. (doi:10.1111/j.1432-1033.2004.04404.x)
  112. Reijenga KA, van Megen YMGA, Kooi BW, Bakker BM, Snoep JL, van Verseveld HW, Westerhoff HV. 2005 Yeast glycolytic oscillations that are not controlled by a single oscillator: a new definition of oscillator strength. *J. Theor. Biol.* **232**, 385–398. (doi:10.1016/j.jtbi.2004.08.019)
  113. Ingalls BP, Sauro HM. 2003 Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories. *J. Theor. Biol.* **222**, 23–36. (doi:10.1016/S0022-5193(03)00011-0)
  114. Wellstead P, Bullinger E, Kalamatianos D, Mason O, Verwoerd M. 2008 The role of control and system theory in systems biology. *Annu. Rev. Control* **32**, 33–47. (doi:10.1016/j.arcontrol.2008.02.001)
  115. Westerhoff HV, Kell DB. 1996 What bio technologists knew all along. . .? *J. Theor. Biol.* **182**, 411–420. (doi:10.1006/jtbi.1996.0181)
  116. Kholodenko BN, Westerhoff HV. 2004 *Metabolic engineering in the post genomic era*. Wyomndham, UK: Horizon Bioscience.
  117. Snoep JL, Yomano LP, Westerhoff HV, Ingram LO. 1995 Protein burden in *Zyomonas-mobilis*—negative flux and growth-control due to overproduction of glycolytic-enzymes. *Microbiology* **141**, 2329–2337. (doi:10.1099/13500872-141-9-2329)
  118. Brockman IM, Prather KLJ. 2015 Dynamic metabolic engineering: new strategies for developing responsive cell factories. *Biotechnol. J.* **10**, 1360–1369. (doi:10.1002/biot.201400422)
  119. Venayak N, Anesiadis N, Cluett WR, Mahadevan R. 2015 Engineering metabolism through dynamic control. *Curr. Opin. Biotechnol.* **34**, 142–152. (doi:10.1016/j.copbio.2014.12.022)
  120. Klumpp S, Zhang ZG, Hwa T. 2009 Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**, 1366–1375. (doi:10.1016/j.cell.2009.12.001)
  121. Weisse AY, Oyarzun DA, Danos V, Swain PS. 2015 Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl Acad. Sci. USA* **112**, E1038–E1047. (doi:10.1073/pnas.1416533112)
  122. Kirk DE. 1970 *Optimal control theory; an introduction*. Englewood Cliffs, NJ: Prentice-Hall.
  123. Yi TM, Huang Y, Simon MI, Doyle J. 2000 Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA* **97**, 4649–4653. (doi:10.1073/pnas.97.9.4649)
  124. El-Samad H, Goff JP, Khammash M. 2002 Calcium homeostasis and parturient hypocalcemia: an integral feedback perspective. *J. Theor. Biol.* **214**, 17–29. (doi:10.1006/jtbi.2001.2422)
  125. Muzzey D, Gomez-Urbe CA, Mettetal JT, van Oudenaarden A. 2009 A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell* **138**, 160–171. (doi:10.1016/j.cell.2009.04.047)
  126. Uygun K, Matthew HWT, Huang YL. 2006 DFBA-LQR: an optimal control approach to flux balance analysis. *Ind. Eng. Chem. Res.* **45**, 8554–8564. (doi:10.1021/le060218f)
  127. Kahn D, Westerhoff HV. 1993 The regulatory strength—how to be precise about regulation and homeostasis. *Acta Biotheor.* **41**, 85–96. (doi:10.1007/Bf00712777)
  128. Bruggeman FJ, Westerhoff HV, Hoek JB, Kholodenko BN. 2002 Modular response analysis of cellular regulatory networks. *J. Theor. Biol.* **218**, 507–520. (doi:10.1006/jtbi.2002.3096)
  129. Klipp E, Heinrich R, Holzhutter HG. 2002 Prediction of temporal gene expression—metabolic optimization by re-distribution of enzyme activities. *Eur. J. Biochem.* **269**, 5406–5413. (doi:10.1046/j.1432-1033.2002.03223.x)
  130. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U. 2004 Just-in-time transcription program in metabolic pathways. *Nat. Genet.* **36**, 486–491. (doi:10.1038/ng1348)
  131. Chin CS, Chubukov V, Jolly ER, DeRisi J, Li H. 2008 Dynamics and design principles of a basic regulatory architecture controlling metabolic pathways. *PLoS Biol.* **6**, 1343–1356. (doi:10.1371/journal.pbio.0060146)
  132. Bartl M, Kotzing M, Schuster S, Li P, Kaleta C. 2013 Dynamic optimization identifies optimal programmes for pathway regulation in prokaryotes. *Nat. Commun.* **4**, 2243. (doi:10.1038/ncomms3243)
  133. Oyarzun DA, Ingalls BP, Middleton RH, Kalamatianos D. 2009 Sequential activation of metabolic pathways: a dynamic optimization approach. *Bull. Math. Biol.* **71**, 1851–1872. (doi:10.1007/s11538-009-9427-5)
  134. Bartl M, Li P, Schuster S. 2010 Modelling the optimal timing in metabolic pathway activation—use of Pontryagin’s Maximum Principle and role of the Golden section. *Biosystems* **101**, 67–77. (doi:10.1016/j.biosystems.2010.04.007)
  135. de Hijas-Liste GM, Klipp E, Balsa-Canto E, Banga JR. 2014 Global dynamic optimization approach to predict activation in metabolic pathways. *BMC Syst. Biol.* **8**, 1. (doi:10.1186/1752-0509-8-1)
  136. Ang J, McMillen DR. 2013 Physical constraints on biological integral control design for homeostasis and sensory adaptation. *Biophys. J.* **104**, 505–515. (doi:10.1016/j.bpj.2012.12.015)
  137. Burrill DR, Boyle PM, Silver PA. 2011 A new approach to an old problem: synthetic biology tools for human disease and metabolism. *Cold Spring Harb. Symp. Quant. Biol.* **76**, 145–154. (doi:10.1101/sqb.2011.76.010686)
  138. Arpino JAJ, Hancock EJ, Anderson J, Barahona M, Stan GBV, Papachristodoulou A, Polizzi K. 2013 Tuning the dials of synthetic biology. *Microbiology* **159**, 1236–1253. (doi:10.1099/Mic.0.067975-0)
  139. Liu D, Xiao Y, Evans BS, Zhang F. 2015 Negative feedback regulation of fatty acid production based on a malonyl-CoA sensor-actuator. *ACS Synth. Biol.* **4**, 132–140. (doi:10.1021/sb400158w)
  140. Xu P, Li LY, Zhang FM, Stephanopoulos G, Koffas M. 2014 Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proc. Natl Acad. Sci. USA*

- 111, 11 299–11 304. (doi:10.1073/pnas.1406401111)
141. Jensen PR, van der Weijden CC, Jensen LB, Westerhoff HV, Snoep JL. 1999 Extensive regulation compromises the extent to which DNA gyrase controls DNA supercoiling and growth rate of *Escherichia coli*. *Eur. J. Biochem.* **266**, 865–877. (doi:10.1046/j.1432-1327.1999.00921.x)
142. Jensen PR *et al.* 1995 Hierarchies in control. *J. Biol. Syst.* **3**, 139–144. (doi:10.1142/S0218339095000137)
143. Rybakova KN, Bruggeman FJ, Tomaszewska A, Mone MJ, Carlberg C, Westerhoff HV. 2015 Multiplex eukaryotic transcription (In) activation: timing, bursting and cycling of a ratchet clock mechanism. *PLoS Comput. Biol.* **11**, e1004236. (doi:10.1371/journal.pcbi.1004236)
144. Bilgin T, Wagner A. 2012 Design constraints on a synthetic metabolism. *PLoS ONE* **7**, e0039903. (doi:10.1371/journal.pone.0039903)
145. Chen Z, Wilmanns M, Zeng AP. 2010 Structural synthetic biotechnology: from molecular structure to predictable design for industrial strain development. *Trends Biotechnol.* **28**, 534–542. (doi:10.1016/j.tibtech.2010.07.004)
146. Goodey NM, Benkovic SJ. 2008 Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **4**, 474–482. (doi:10.1038/Nchembio.98)
147. Chen Z, Rappert S, Zeng AP. 2013 Rational design of allosteric regulation of homoserine dehydrogenase by a non-natural inhibitor L-lysine. *ACS Synth. Biol.* **4**, 126–131. (doi:10.1021/sb400133g)
148. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KLJ, Keasling JD. 2009 Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27**, 753–759. (doi:10.1038/nbt.1557)
149. Mendes P, Kell DB, Westerhoff HV. 1996 Why and when channelling can decrease pool size at constant net flux in a simple dynamic channel. *Biochim. Biophys. Acta* **1289**, 175–186. (doi:10.1016/0304-4165(95)00152-2)
150. Zhang YHP. 2011 Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **29**, 715–725. (doi:10.1016/j.biotechadv.2011.05.020)
151. You C, Zhang YHP. 2013 Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. *ACS Synth. Biol.* **2**, 102–110. (doi:10.1021/sb300068g)
152. Kholodenko BN, Cascante M, Westerhoff HV. 1995 Control-theory of metabolic channeling. *Mol. Cell. Biochem.* **143**, 151–168. (doi:10.1007/bf01816949)
153. Sauro HM, Kholodenko BN. 2004 Quantitative analysis of signaling networks. *Prog. Biophys. Mol. Biol.* **86**, 5–43. (doi:10.1016/j.pbiomolbio.2004.03.002)
154. McClean MN, Mody A, Broach JR, Ramanathan S. 2007 Cross-talk and decision making in MAP kinase pathways. *Nat. Genet.* **39**, 409–414. (doi:10.1038/Ng1957)
155. Mather WH, Hasty J, Tsimring LS, Williams RJ. 2013 Translational cross talk in gene networks. *Biophys. J.* **104**, 2564–2572. (doi:10.1016/j.bpj.2013.04.049)
156. Kolodkin A, Simeonidis E, Westerhoff HV. 2013 Computing life: add logos to biology and bios to physics. *Prog. Biophys. Mol. Biol.* **111**, 69–74. (doi:10.1016/j.pbiomolbio.2012.10.003)
157. Dolan J, Anderson J, Papachristodoulou A. 2012 A loop shaping approach for designing biological circuits. In *2012 IEEE 51st Annual Conf. on Decision and Control (CDC), Maui, HI, USA, 10–13 December 2012*, pp. 3614–3619. (doi:10.1109/CDC.2012.6426405)
158. Del Vecchio D. 2013 A control theoretic framework for modular analysis and design of biomolecular networks. *Annu. Rev. Control* **37**, 333–345. (doi:10.1016/j.arcontrol.2013.09.011)
159. Raser JM, O'Shea EK. 2005 Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010–2013. (doi:10.1126/science.1105891)
160. Raj A, van Oudenaarden A. 2008 Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226. (doi:10.1016/j.cell.2008.09.050)
161. Chalancon G, Ravarani CNJ, Balaji S, Martinez-Arias A, Aravind L, Jothi R, Babu MM. 2012 Interplay between gene expression noise and regulatory network architecture. *Trends Genet.* **28**, 221–232. (doi:10.1016/j.tig.2012.01.006)
162. Kim KH, Sauro HM. 2011 Measuring retroactivity from noise in gene regulatory networks. *Biophys. J.* **100**, 1167–1177. (doi:10.1016/j.bpj.2010.12.3737)
163. Levine E, Hwa T. 2007 Stochastic fluctuations in metabolic pathways. *Proc. Natl Acad. Sci. USA* **104**, 9224–9229. (doi:10.1073/pnas.0610987104)
164. Puchalka J, Kierzek AM. 2004 Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks. *Biophys. J.* **86**, 1357–1372. (doi:10.1016/S0006-3495(04)74207-1)
165. Zamora-Chimal C, Santillan M, Rodriguez-Gonzalez J. 2012 Influence of the feedback loops in the trp operon of *B. subtilis* on the system dynamic response and noise amplitude. *J. Theor. Biol.* **310**, 119–131. (doi:10.1016/j.jtbi.2012.06.014)
166. Rocco A. 2009 Stochastic control of metabolic pathways. *Phys. Biol.* **6**, 016002. (doi:10.1088/1478-3975/6/1/016002)
167. Kim KH, Sauro HM. 2010 Sensitivity summation theorems for stochastic biochemical reaction systems. *Math. Biosci.* **226**, 109–119. (doi:10.1016/j.mbs.2010.04.004)
168. Oyarzun DA, Lugagne JB, Stan GB. 2014 Noise propagation in synthetic gene circuits for metabolic control. *ACS Synth. Biol.* **4**, 116–125. (doi:10.1021/sb400126a)
169. Zavlanos MM, Julius AA. 2011 Robust flux balance analysis of metabolic networks. In *2011 American Control Conf., San Francisco, CA, USA, 29 June–1 July 2011*, pp. 2915–2920.
170. Bates DG, Cosentino C. 2011 Validation and invalidation of systems biology models using robustness analysis. *IET Syst. Biol.* **5**, 229–244. (doi:10.1049/iet-syb.2010.0072)